TNO Defence, Security and Safety

TNO report

TNO-DV3 2005 A245

The role of operator state assessment in adaptive automation

Date	January 2006
Author(s)	J.A. Veltman, M.Sc.
	C. Jansen. M.Sc.

Classification report	Ongerubriceerd
Classified by	KLTZ J.J. Teertstra
Classification date	December 2005
	(This classification will not change)
Title	Ongerubriceerd
Managementuittreksel	Ongerubriceerd
Abstract	Ongerubriceerd
Report text	Ongerubriceerd
Appendices	Ongerubriceerd

8

|	
Copy no	
No. of copies	17
Number of pages	52 (incl. appendices, excl. RDP & distributionlist)
Number of appendices	3

AQ F06-07-6404

The role of operator state assessment in adaptive automation

Een vorm van adaptieve automatisering waarbij de taakbelasting automatisch wordt afgestemd op de toestand van de operator wordt in de literatuur gezien als een veelbelovende verbetering in de mens-systeeminteractie. De toestand kan bepaald worden met behulp van fysiologische maten. In dit rapport wordt beargumenteerd dat deze manier van adaptieve automatisering niet altijd wenselijk is. Bovendien laten experimentele gegevens zien dat het nauwkeurig meten van de toestand van een operator zeer lastig is.

Operator voert een taak uit waarbij fysiologische gegevens worden gemeten om een inschatting te krijgen van zijn mentale werkbelasting.
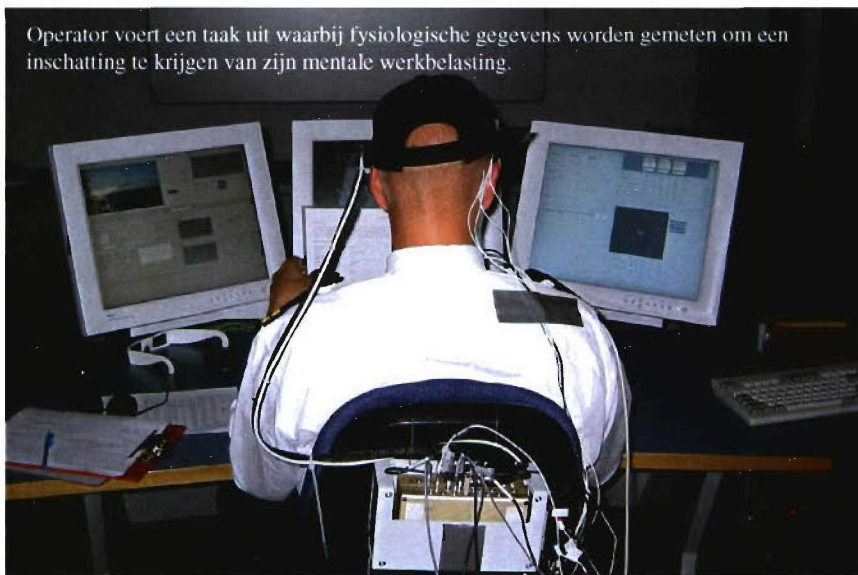
Beschrijving van de werkzaamheden

Over het gebruik van fysiologische maten voor mentale werkbelasting en de toepassing bij adaptieve automatisering is literatuur doorgenomen en gerapporteerd. Tevens is een model opgesteld aan de hand waarvan de relatie tussen de toestand van een operator diens informatieverwerking en de interactie met de taakomgeving kan worden beschreven.
Daarnaast is een experiment uitgevoerd waarbij de toestand van operators is bepaald aan de hand van fysiologische reacties. Hierbij is gekeken naar veranderingen in fysiologische reacties tijdens de taakuitvoering en de relatie met de moeilijkheid van de taak.

Probleemstelling

Het is steeds beter mogelijk om de mate van automatisering te variëren tijdens taakuit- voering (adaptieve automatisering). Hierbij zou gebruik gemaakt kunnen worden van informatie over de toestand van een operator. Het achterliggende idee hierbij is dat de taakmoeilijkheid wordt verlaagd als de operator mentaal zwaar belast is. Hiervoor zouden fysiologische reacties gebruikt

kunnen worden waaruit de mentale belasting wordt afgeleid.
In opdracht van de Koninklijke Marine heeft TNO Defensie en Veiligheid een verkennend onderzoek uitgevoerd naar de mogelijkheden om de toestand van een operator te betrekken bij adaptieve automatisering.

Resultaten en conclusies

Uit de literatuur blijkt dat er goede resul- taten verwacht worden van het gebruik van fysiologische maten bij adaptieve automatisering. Echter, op basis van het model dat is opgesteld kan hier een aantal kanttekeningen bij worden geplaatst. Een belangrijk aandachtspunt hierbij is het adaptieve gedrag van een operator. Deze past zich continu aan taakeisen aan. Indien

de taakomgeving zich ook aanpast aan de operator is de kans op een instabiel mens-machinesysteem groot.

De fysiologische metingen tijdens het experiment laten bovendien zien dat het inschatten van de mentale werkbelasting gedurende de taakuitvoering nog niet goed mogelijk is met de gebruikte maten. Wel is het mogelijk om fysiologische maten te gebruiken voor een inschatting van de werkbelasting gedurende langere tijdsperioden, maar dit is onvoldoende voor het gebruik bij adaptieve automatisering.

Toepasbaarheid

Deze resultaten geven een beter inzicht in de complexe relatie tussen taakbelasting, prestatie en de toestand van een operator. Fysiologische maten kunnen goed gebruikt worden om een algemeen beeld te krijgen van de werkbelasting, maar zijn vooralsnog niet geschikt om te gebruiken bij adaptieve automatisering.

PROGRAMMA	PROJECT
Programmabegeleider Kltz J.J. Teertstra	Projectbegeleider Kltz J.J. Teertstra
Programmaleider Drs. J.H. van Delft, TNO Defensie en Veiligheid	Projectleider Drs. J.A. Veltman, TNO Defensie en Veiligheid
Programmatitel Menselijke supervisie en geavanceerde taakuitvoering	Projecttitel Operator state assessment
Programmanummer V055	Projectnummer 013.63030
Programmaplanning Start 1-1-2003 Gereed 31-12-2005	Projectplanning Start 1-1-2003 Gereed 31-12-2005
Toezichthouder -	
Frequentie van overleg	Projectteam drs. J.A. Veltman, drs. C. Jansen, A. Wennemers

Samenvatting

Computersystemen zijn in staat om steeds meer taken van menselijke operators over te nemen. Het is echter niet altijd wenselijk om de operator belangrijke taken uit handen te nemen. Het adaptief maken van de automatisering waarbij de mate van automatisering wordt afgestemd op de omstandigheden zou een grote verbetering betekenen voor de prestatie van het mens-machine systeem. Eén van de parameters waarop de automatisering kan worden afgestemd is de toestand van de operator. Hierbij is het idee dat de computer taken van een operator overneemt of minder informatie aanbiedt op het moment dat de operator te maken heeft met een hoge mentale werkbelasting.

Dit rapport is opgebouwd uit drie delen. In het eerste deel wordt een overzicht gegeven van literatuur over fysiologische maten voor mentale belasting en het gebruik hiervan bij adaptieve automatisering. Hieruit blijkt dat van fysiologische werkbelastingsmaten een belangrijke bijdrage wordt verwacht bij toekomstige adaptieve systemen.

In het tweede deel van het rapport wordt een model gepresenteerd aan de hand waarvan de relatie tussen de toestand van een operator, de menselijke informatieverwerking en de interactie met de buitenwereld wordt uitgelegd. We betogen dat menselijk gedrag zich kenmerkt door een continue aanpassing aan de buitenwereld. De belangrijkste middelen voor dit adaptieve gedrag zijn het veranderen van de toestand (bijvoorbeeld meer inspanning leveren als de taakomgeving meer eist), of het verlagen van de taak-doelen (waardoor de prestatie omlaag gaat). Op basis van dit model wordt beargumenteerd dat een adaptief systeem waarbij de mate van automatisering direct gekoppeld wordt aan de toestand van een operator niet per definitie goed zal werken. Dit komt omdat de operator zich aan de taak aanpast terwijl de taak zich aan de operator probeert aan te passen.

In het derde deel wordt een experiment beschreven waarbij een aantal fysiologische maten zijn gemeten bij operators van de Koninklijke Marine tijdens het uitvoeren van drie verschillende missies in een gesimuleerde taakomgeving. Gemeten zijn hartslag-frequentie, hartslagvariabiliteit, ademhaling en oogknipperingen.

Er is onder andere gekeken naar de fysiologische reacties tijdens de missies. Dit is belangrijk voor het gebruik van adaptieve automatisering omdat de toestand nauw-keurig gevolgd moet kunnen worden om de taakeisen te kunnen aanpassen. De resultaten lieten zien dat de fysiologische reacties niet systematisch varieerden als functie van de moeilijkheid van de taak. Uit andere gegevens is gebleken dat de variatie in moeilijkheid gedurende de missie ook niet groot was, waardoor systematische veranderingen in de fysiologische reacties mogelijk niet meetbaar waren.

Als alle resultaten worden betrokken concluderen we dat het niet waarschijnlijk is dat de onderzochte maten in de toekomst bruikbaar zijn bij adaptieve automatisering.

Summary

Computer systems are capable to take over more and more tasks from human operators. However, it is not always desirable to take away important tasks from operators. Automation of which the level is made dependent on the situation may be an improvement for the performance of the human-machine system. One of the possible parameters that can be used for this so-called adaptive automation is the state of the operator. The idea is that the computer can take over tasks or provides less information when the mental workload of the operator is high.

This report contains three main parts. The first part provides an overview of the literature on physiological measures for mental workload and the use of these measures in adaptive automation. The literature shows that the use of physiological measures as a parameter for adaptive automation is regarded as very promising.

The second part provides a model that is used to describe the relation between the state of the operator, the human information processing and the interaction with the outside world. It is argued that human performance is characterized by a constant adaptation to the outside world. Important means for this adaptive behavior are the change of the state (e.g., by investing more mental effort when tasks becomes more difficult) or lowering the task goals (accepting a lower level of performance). Based on this model we argue that an adaptive system in which the level of automation is directly related to the state of the operator is not likely to function. This is because there are in fact two adaptive systems, the operator is adapting to changing task demands and the computer is adapting to changes in the state of the operator.

The third part describes an experiment in which several physiological measures were monitored during the task performance of operators from the Netherlands Navy. These measures were heart rate frequency, heart rate variability, respiration, and eye blinks. The analysis of the data was focused on changes in physiological reactions during the missions and the relations with changing task demands. This is relevant for the application of physiological measures as a parameter of adaptive automation, because the state of the operator has to be estimated accurately for this application.

The results show that changes in the physiological measures were not related to changes in task demand. Furthermore, the measures did not show a congruent pattern of results. The changes in task demand appeared to be moderate and therefore, systematic changes in physiological reactions were not likely to be present. Taking all results into account it can be concluded that the physiological measures that are used in the present experiment are not likely to be useful for adaptive automation in the future.

Contents

1 Introduction

Many years ago systems started to take over simple and routine tasks from operators. With the latest techniques, more complex tasks can be taken over by systems. Everyone can see this evolution in automation in cars. In the fifties, the car driver had to wait before he could switch off the wipers until the wipers were in the lower position (otherwise, they would stop at the middle of the windscreen). Nowadays, nobody would think about this when they switch off the wipers. A newer automation system that is almost common in most cars is the cruise control. The driver does not have to control the speed any more when this system is turned on. In the newest version, the car will even automatically slow down when it approaches a car in front of the own car. New techniques often also increase the number of tasks to be performed. Examples are the in-car telephones and information systems.

In control rooms on board frigates a similar evolution in automation and growth of information systems can be observed. The result is that the complexity of operations that can be conducted by frigates has strongly increased. Moreover, a reduction of the number of operators is going on which also results in more tasks for the remaining operators. For most tasks, the operator is the one who is still in control. The operator has to make a decision about what actions have to be taken. Even if most tasks are taken over by the systems, the operator has to build up adequate situation awareness. He must be able to perceive the relevant information, to understand the content, and be aware of the consequences for the near future in order to anticipate to changes in the environment [Endsley, 1995]. For adequate situation awareness, the operator must have an active role but he must not be overloaded with information. It would therefore be very good if a system would take into account the state of the operator. If the system presents information to the operator when he is overloaded, then it is likely that the newest information will be missed by the operator or that the operator will skip other relevant tasks.

This requires a new way of automation. The level of automation should dynamically change with changing circumstances. This is called adaptive automation. The following components are important for adaptive information:

1   a model of the context in which the tasks have to be performed;
2   a model of the system;
3   a model of the tasks (e.g. an estimation of the level of demand the tasks put on an operator);
4   a model of the operator state (e.g. the workload of the operator).

Physiological measures have been used often for the assessment of the operator state, especially to measure the amount of mental effort that the operator invests into the task to cope with the task demands. This report emphasises the role of state assessment in adaptive automation. A literature overview is presented about the relation between physiological measures and adaptive automation and a model is presented that describes the relation between operator state and information from a system. Furthermore, results of and experiment are presented in which the state of operators is assessed with several physiological measures.

2   Operator state

2.1   Relation between mental workload and performance

Figure 1 describes the relation between mental workload and operator performance. This relation is not straightforward. When the workload is low (low task demands) it is easy to have an optimal level of performance. However, when the task requires little attention, human beings have difficulty to remain alert. So, when new information is presented, operators are likely to miss it and the performance can decrease considerably. In a normal workload situation, the operator is interacting with the system on a regular basis in which it is not difficult to maintain an optimal level of performance. In a high workload situation operators are only able to maintain a high level of performance when they exert additional effort. This can not be prolonged for a long time without costs. Operators become more fatigued, which might result in more errors. Furthermore, the recovery time after the work period will be longer. In an overload situation (too much information), the operator can not get an acceptable level of performance any more. It is likely that the operator will even stop to exert additional effort, because this will not help him anymore.

The relation between task demand and performance is different for individuals due to differences in talent and level of training. This relation can also change in different circumstances within one operator due to a state change. When the state changes, the task demand that results in a normal workload might result in a high workload situation when the state is not optimal. State changes can have many different causes such as working at unusual times (night work), fatigue, sickness and external stressors (e.g., loud noises, vibration, ship movements).
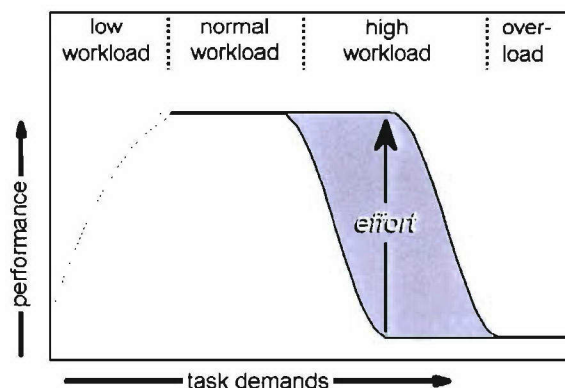
Figure 1   Relation between task demand and operator performance.

2.2   Model of operator state

We developed a model (see Figure 2) that describes the relation between operator state and information processing. This model includes a loop that describes the different information processing stages (perception, central processing and action selection) and a loop that describes the state regulation. Task goals play a central role in the model in a

way that it drives the intensity of the information processing. Task goals (the level of performance that the operator wants to achieve) are compared with the perceived performance. If the level of performance is too low, the intensity of the information-processing loop has to increase. This is often only possible if the state of the operator makes a higher intensity of the information processing possible. In the lower loop, the required state is compared with the actual state. If these two do not match, the state can be improved by exerting additional effort. An alternative to increasing mental effort is changing the task goals (e.g., allow more errors or skip less relevant tasks). It becomes more likely that the perceived performance matches the required performance, which result in a new equilibrium.

Task performance can be seen as an adaptive process to changing situations in the environment. The present model describes this adaptive behaviour of an operator.
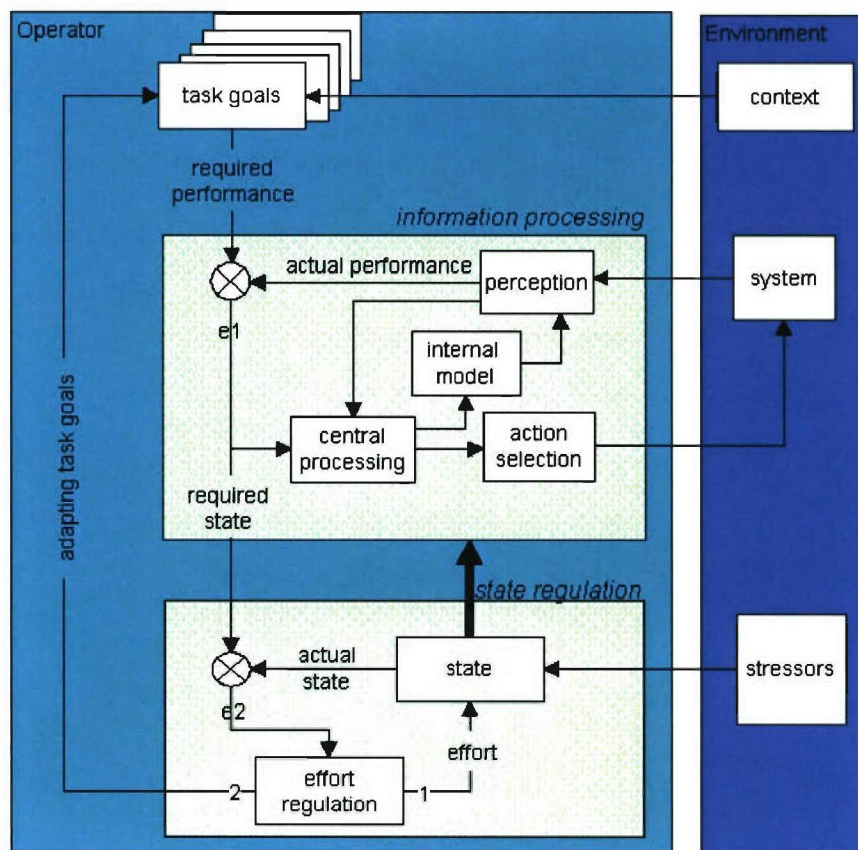

Figure 2   Operator state model (see text for explanation).

Carver and Scheier (1998) showed that human behaviour is strongly driven by differences between internal goals and sensory information. This idea is based on the perceptual control theory [PCT, Powers, 1973]. Contrary to many human information processing theories that assume that behaviour is mainly driven by information input, the PCT assumes that behaviour is mainly driven by goals. Behaviour is meant to change the incoming information in order to reduce the differences between goals and sensory information.

The self-regulation of human behaviour is comparable to many man-made control systems, such as a cruise control, which regulates the speed by comparing the goal (required speed) with sensory information (actual speed). The system can control sensory information (speed) by changing the power of the engine.

Obviously, a human being does have many more goals and many more sensors to control but basically PCT assumes the mechanism not to be much different. There are different hierarchical loops (probably up to nine) involved in the control of sensory information. The output of a higher order control loop can be a goal of a lower order control loop. For example, if the goal of an operator is to perform very good during a mentally demanding task, the physiological system has to adapt to this situation by increasing the blood flow in the brains. This output of this higher order control system results in a *goal* for a lower control system (i.e. a new blood pressure reference value for the cardiovascular control system). There are several blood pressure sensors in the veins and the cardiovascular control system can change this sensory information by increasing the heart rate, increasing the cardiac output, decreasing the amount of blood in the extremities of the body etc. The model in Figure 2 only has two hierarchical loops of which the required state is an output of the highest loop and a goal for the lowest loop.

The PCT is also used in mental workload models of Hockey (2003) and Hendy, East and Farrel (2001). The model of Hockey mainly uses PCT to describe state regulation, whereas Hendy et al. use the PCT to describe information processing. The model in Figure 2 is a combination of these models.

A general information-processing loop is included to describe the stages of information processing of an operator dealing with a system (perception, central processing and action selection). Information to be processed can come from the systems that the operators have to deal with, or from an internal model that we built up from the environment around us.

The perceived information, and in particular, the perceived performance is compared with the task goals. The intensity of the information-processing loop is adjusted depending on the difference between the required and actual performance.

The state of the operator is important for the information processing. A non-optimal state will negatively affect the information processing. For example, it is commonly known that performing a complex task just after awakening or after 10 hours of intensive work will be difficult. The state has to be adjusted when the required state does not match the actual state (high 'e2' in the model).

One way to deal with this increased state requirement is to invest more mental effort (route '1' from the 'effort regulation' box). However, effort investment does involve direct costs (such as fatigue) and indirectly by negative feelings that makes further effort investment more difficult. Another mechanism to reduce the difference between the required and the actual state is to adapt the task goals (route '2' from the 'effort regulation' box). When the required performance decreases, for example by accepting a lower work rate and/or accepting more errors, the information processing will be less intense. This reduces the required state, which will result into a balance in the lower control loop.

The *context* is an important variable in the model. It affects the priorities of the task goals and the consequences of changing the task goals. For example, making errors is not an option during an examination or a war situation, while this does not have many consequences during a normal training. Therefore, the state of the operator is likely to be different during a war or an examination compared to a training situation, even when the same task is performed.

In many situations, the task goals are just one set of goals among many other goals such as keeping rest, going to a toilet, have a conversation, going away for a cigarette, etc. The context is important for keeping the task goals the primary ones.

The goal of the highest level is the task goal. This can be interpreted as the intentions of the operator.

Stressors in this model refer to environmental factors that affect the state of the operator such as extreme temperatures, vibration, G-forces.

The model illustrates that there is no direct relation between physiological measures that are used as 'state' estimators and information load. In other words, making a task more difficult will not automatically result in changes in physiological reactions. Some examples might illustrate this relation.

- When the operator is very well trained, more information does often not result in a more 'intensive' information processing and therefore, will not result in a higher 'required state'.
- An increase in information load may result in an adaptation of the task goals. For example the operator can take more time to perform the task, skip some tasks, or will be satisfied with more errors.
- When there is a discrepancy between the information from the system and the internal model, then the intensity of the information processing (and the 'required state') may increase considerably. For example, when information about the position of an aircraft is different from earlier observations (or does not match to the internal model), many extra checks have to be performed in order to validate this information and/or the earlier information. This will mostly result in a higher state [Veltman & Gaillard, 1999]. When the same information does match to the internal model, then no effects on state are expected.

The model can also be used to understand differences between subjective and physiological workload measures. Earlier experiments showed that subjective workload measures are very sensitive to increases in the error signal (e1 and e2), whereas physiological measures are more sensitive to state changes [Veltman & Gaillard, 2003].

3   Literature review of adaptive automation

Adaptive automation refers to systems in which both the user and the system can initiate changes in the level of automation. Such systems can be described as either adaptable or adaptive. In *adaptable systems* the user actively initiates the presentation mode of information or the allocation of functions between the operator and the system. In *adaptive systems*, the system can initiate the presentation mode or the allocation of functions.
The state of the operator can be an important parameter for adaptive systems. The 'operator state' is here restricted to the amount of operator's mental workload during task execution. We will start with a short description of problems related to static automation (i.e., task allocation does not change during performance) to illustrate the expected benefits of psychophysiology-based adaptive automation. One should keep in mind, however, that very often performance benefits from (static) automation. Here, we focus on those instances in which that is not the case, to gain insight into how to improve human-automation co-operation to optimise task performance.

From the beginning, automation has been applied to tasks that were too dull, dirty, or dangerous for human operators to perform. Inspired by the success of these applications and facilitated (and pushed) by technological progress, more and more kinds of tasks were automated, including tasks that require intelligence. There are quite some advantages of automation, including increase in safety, reliability, and precision, and a decrease in operator workload [Wiener & Curry, 1980]. However, automation did not always lead to performance improvement. Apparently, some tasks are performed better by human operators than by machines. More than 50 years ago, Fitts presented the MABA-MABA list ('Men Are Better At' and 'Machines Are Better At'), stating rules-of-thumb on what (not) to automate [Fitts, 1951]. For example, men are better at reasoning inductively and at perceiving patterns of stimuli, whereas machines are superior in reasoning deductively and in responding quickly to control signals. Such a list can be helpful, but who should do a particular task that requires deductive reasoning on patterns of stimuli? One of the criticisms on Fitts' method is that it does not make clear how to resolve such conflicts. Furthermore, the method emphasises (sub)task allocation to either men or machine without addressing the issue of integrating men's and machine's efforts in their co-operation in getting the job done [Hoc, 200]. It seems to be impossible to make a short list of generally applicable, <u>decisive</u> rules on what (not) to automate. Rather, experts' knowledge on success and failure of automation in similar situations seems to guide the design of human-automation co-operation [Endsley, 1996; Parasuraman & Mouloua, 1996; Wiener et al., 1980; Parasuraman & Riley, 1997]. In this process, knowledge on the consequences of partly automated tasks on human operator performance is most relevant, especially when the human operator is responsible for the overall performance of human-machine co-operation. In the next paragraph we provide a brief overview of human operator errors that can possibly result from static automation.

3.1   Human operator error associated with static automation

Bainbridge's (1983) message seemed to be that human operators are impressive problem solvers as long as there is no time pressure, thereby implying that such tasks should never be automated. In addition, she stated that static automation cannot compensate for the inadequate human performance under time pressure. She clearly

summarised the problems associated with static automation as follows: 'By taking away the easy parts of his task, automation can make the difficult parts of the human operator's task more difficult' [Bainbridge, 1983]. To elaborate, due to automation, the task may become more difficult and hence performance more erroneous because the human operator is not actively involved in every aspect of the task. The resulting suboptimal situational awareness can lead to sub-optimal human performance on the non-automated aspects of the tasks. Automation can also increase the operator's mental workload, which flies in the face of the designer's intentions: Sometimes, automation increases workload, because the role of an operator shifts from an active participant to a passive monitor. This monitoring role brings about another kind of workload, that is sometimes even higher [Sarter & Woods, 1995; Kirlik, 1993]. Another source of error is that the human operator does not train the manual skills to immediately intervene when the machine is unable to perform adequately. Even worse, the human operator monitoring the automated tasks can be less likely to detect such machine malfunctioning [Parasuraman et al., 1996; Wickens & Kessel, 1979], especially when other tasks need to be performed [Parasuraman, Molloy, & Singh, 1993]. Another cause of sub-optimal monitoring performance may be that the operator without manual experience on the automated task relies on different (less effective?) cues [Kessel & Wickens, 1982].

3.2   Intermediate human control

A promising alternative to continuous static automation is to have the operator perform the task manually at stated intervals [Bainbridge, 1983]. Performance may improve because such intermediate manual control could take away operator's boredom, vigilance decrement and underload. However, Parasuraman, Molloy, and Singh (1993) concluded that these factors cannot account for complacency. Farrell and Lewandowsky (2000) state that complacency is due to the operator's inability to suddenly switch to a different cognitive operation. For example, when the operator needs to intervene because of an automation failure, he does not have the task-appropriate response available at once. Their connectionist model, based on this idea, was able to account for several complacency-related phenomena, among which the benefit of intermittent manual control. Also, experimental results show improved performance for intermediate manual control. Kessel and Wickens (1982) observed enhanced performance in monitoring the dynamics of a 2-dimensional pursuit display when participants had prior manual experience. Parasuraman, Mouloua, and Molloy (1996) reported similar findings in a more complex task setting: Performance of monitoring an automated engine status task while simultaneously performing a tracking and a fuel management task was improved after a brief period of manually performing that engine status task. This positive effect of intermediate manual control on monitoring performance sustains for a longer period of time [Mouloua, Parasuraman, & Molloy, 1993]. Compared to continuous static automation, automation with intermediate manual control will increase mental workload (of course during manual control, but also during the transition between control modes) but will also increase situation awareness. For finding the right balance in the trade-off between workload and situation awareness, at least two issues need to be resolved. First, how frequent may a switch between automatic and manual control occur? Epharth and Young (1981) reported that operators need some time to adjust after switching between activity modes (e.g., monitoring to controlling), even if the switch is initiated by themselves. This may account for poorer performance with excessively frequent cyclings between manual control and full automation [Hilburn et al., 1993] and with extremely short cycle duration [Scallen et al.,

1995]. But how frequent is too frequent, and how short is too short? Second, what or who should initiate the switch, and on what grounds? This will be discussed in the next section.

3.3 Adaptive automation

With intermediate manual control, task allocation to either man or machine is dynamic, resulting in a variable degree of automation during task execution. This form of automation is referred to as adaptive automation. It is important to note that adaptive automation should be considered as being radically different from static automation: whereas continuous static automation is working *for* the operator, adaptive automation should be seen as an interactive aid working *with* the operator (i.e., man-machine co-operation). A clever method that controls the task allocation in adaptive automation (i.e., that determines when what (not) to automate) may very well be the technological ingenuity. Bainbridge (1983) referred to as needed for countering the inadequacy of static automation. Several approaches for developing such a method can be thought of [Morrison & Gluckman, 1994; Scerbo, 1996]. A major distinction between methods is whether the operator or the computer initiates a switch between manual and automation modes [Hilburn et al., 1993]. Harris, Hancock, Arthur and Caird (1991) found support for initiation by the operator: Participants performed better on resource management task when they had control over invoking automation on a compensatory tracking task, as compared to conditions in which either the tracking task was never automated or it was always automated. A pitfall for such a method could be that humans generally underestimate their capabilities, even for physical tasks [Holding, 1983]. As a result, the operator could switch to automation too soon and too often, thereby impairing his situational awareness.

Dynamic task allocation initiated by a computer can be based on different methods. Switches between manual and automatic control could occur at, for example, predefined (e.g., every 20 minutes) or at random instances. A major drawback of these methods is that the switch may be poorly timed (like half-way task execution), resulting in performance degradation. More clever methods could reallocate tasks based on (a) the changes in operator performance (e.g., automating more Air Traffic Control subtasks in response to more operator errors), (b) the changes in task complexity (e.g., automating more Air Traffic Control subtasks in response to more planes), or (c) the changes in operator functional state (e.g., automating more Air Traffic Control subtasks in response to an operator's unacceptably high mental workload).

Methods based on (a) have been found to be successful. For example, Kaber and Riley (1999) reported that performance degradation on a secondary task may herald degradation on the primary task. However, this method has as an important drawback that one is chasing the facts: the to be prevented performance degradation has already occurred. In addition, this method is less suitable when overt operator performance is sparse, as in automated systems. Methods based on changes in task complexity (b) could in principle prevent high workload and performance degradation because tasks are automated before the operator is confronted with their increased complexity. However, if the operator's workload was not unacceptably high at the occurrence of the increase in task complexity, it may be better not to automate for keeping the operator's situational awareness as high as possible. Methods based on (c) attempt to tune task allocation to the preparedness of the operator for performing the tasks. A successful implementation of such a method would guarantee the right balance in the trade-off between operator mental workload and operator situational awareness. We agree with

many others [Byrne & Parasuraman, 1996], that this is a promising method because it aims to predict, and possibly prevent, performance degradation in man-machine co-operation. For example, Parasuraman, Mouloua, and Hilburn (1999) have reported performance benefits when task reallocation is closely matched with operator workload, providing evidence for the need for an optimal coupling of automation level and level of operator workload in adaptive automation [Parasuraman et al., 1992]. The current study aims to investigate the feasibility of using psychophysiological measures for adaptive automation.

3.4 Assessing operator Functional state and adaptive automation

In order to be able to use psychophysiological measures to control adaptive automation, one should be sure these measures correctly assess the operator state. Psychophysiological measures have already been used for monitoring the state of operators and pilots in order to prevent (sudden and drastic) degradation in performance. For example, induced loss of consciousness due to high G-forces (+Gz) can be detected by EEG and ECG [Whinnery, Glaister, & Burton, 1987], and spontaneous variability in palmar skin conductance can trigger an auditory alarm for alerting an operator [Yamamoto & Isshiki, 1992; Satchell, 1993]. Applications that use psychophysiological measures have been successful in assessing major changes in operator state, as falling asleep or fainting away. However, sometimes performance degradation results from more subtle changes in the state of the operator. For example, mental overload could result in incorrect decisions at crucial moments; complacency as a result of mental underload could even result in not perceiving the 'cruciality' of such moments.

Measures to be used in adaptive automation should be able to differentiate between several subtly different operator states. An important contribution to this endeavour was made by Wilson (1993) who demonstrated in an offline analysis that psychophysio-logical measures can differentiate peak levels in mental workload at different phases of a flight mission for different crew members. In a latter study, Wilson and colleagues successfully used an artificial neural network to classify operator functional state in a complex task setting with 4 tasks [Wilson, Lambert, & Russell, 2000]. EEG (delta, theta, alpha, and beta bands), ECG inter beat intervals, EOG blink rates and blink closure durations, and respiration rates were recorded in three conditions: baseline recording, low workload, high workload. In an off-line analysis, participant-specific-trained neural networks were able to correctly classify psycho-physiological states to either of the three conditions in 98.5% of the cases. (They also used the trained neural networks in an adaptive automation setting, as will be discussed in a later section). Notwithstanding this success, predicting and preventing performance degradation using psychophysiological measures turns out to be very challenging. Hockey, Gaillard, and Burov have recently published an excellent overview of the state of the art on operator functional state assessment, with contributions of experts in the field [Hockey, Gaillard, & Burov, 2003]. It deals on theoretical and methodological frameworks, methods of assessment, and contains a section devoted to the application of operator state assessment in adaptive automation.

[Scerbo et al., 2001] provide an overview of psychological measures that may qualify for adaptive automation: eye blinks, respiration, cardiovascular activity, speech. They conclude that cardiovascular measures (heart rate, heart period, heart rate variability) are most suitable because they are reliable, easy to record, and minimally intrusive.

Slightly less easy to record and more intrusive, but not less promising appears to be the use of cortical measures. The same group of researchers has quite some experience with adaptive automation based on EEG measures (delta, theta, alpha, beta bandwidths).

Heart rate variability (HRV) often shows greater efficacy in detecting gross changes in workload rather than refined gradations [Jorna, 1992]. It has the advantage of being truly non-intrusive. It may reflect a mixture of cognitive processing demands and energetic processes (i.e., compensatory efforts). Byrne, Chun, Hilburn, Molloy & Parasuraman (1994) report that in contrast with earlier findings with single task conditions, in a multitask environment, only group decreases in HRV in response to task load are observed, and no relationship to individual differences in subjective ratings of effort seems to emerge. Wilson, Fullenkamp and Davis (1994) found correlations between subjective measure for task difficulty and evoked cortical potentials, heart rate, and blink rate respiration rate.

Prinzel, Freeman, Scerbo, Milkulka, & Pope (1998) looked at Event Related Potentials (ERP) as another psychophysiological measure for adaptive aiding. This is a change in the electro encephalogram (EEG) after a specific event. Several components in the ERP signal can be distinguished based on its time of occurrence and position on the scalp. The best known component is the P300, which refers to a positive peak in the EEG signal around 300 ms after an event. Kramer, Trejo, and Humphrey (1996) mention that most studies showing the relationship between the P300 component and perceptual/cognitive processing demands were on ERPs elicited on the secondary task; a method that is therefore intrusive. The primary task technique, where ERPs are elicited by discrete events within the task of interest, has the disadvantage that it does not inform about the operator state in between two events. An alternative to these is the irrelevant probe technique [Papanicolaou & Johnstone, 1984]. Normally ERPs are averaged over several repetitions of the stimulus to enhance signal-to-noise ratio, which is inappropriate for moment-to-moment monitoring of operator state. Advantage of ERP is that it can discriminate between different components of mental workload. For example, the P100 component is sensitive to the attention allocation to a particular region of the visual field., the P200 to a particular stimulus match to a predefined template, P300 reflects the stimulus evaluation process, N400 reflects the detection of a semantic mismatch. They conclude that ERP in particular might have some utility as measures of momentary fluctuations in workload and therefore might serve as a trigger for adaptive aiding. In another experiment with 10 Navy radar operators, they used the irrelevant probe procedure with one high probability tone and two low-probability tones. Low-probability tones often elicit larger amplitude ERP than high-probability tones, which facilitates detecting of task dependent changes. Apart from N100, N200, and P300 components, the mismatch negativity (MMN) was of interest, which is best dissociated from other ERP components when elicited by high- and low-probability events that do not require an overt response. They found that N100, N200, and MMN decreased in amplitude with the introduction of the monitoring task as well as with an increase in difficulty. They believe that ERPs elicited by task irrelevant probes can provide a non intrusive method for the assessment of variations in mental workload, but that unfortunately the N100, N200 and MMN components are relatively small in comparison to the ongoing EEG activity, so it is unlikely that they can provide real-time assessment of mental workload. Even though workload may be difficult to assess online, ERP could track attention allocation: Farwell and Donchin (1988) report on the development of an ERP-based communication device in which the P300 component is used to index operators' attention to particular objects in a 6 by 6 matrix of letters and

numbers. Then, peripheral attention could be followed while the operator is fixated to a location on the central visual field. Another application could be the use of the error related negativity (ERN) to monitor if the operator was aware of the error he made [Gehring et al., 1993].

Apart from the ERPs, other information can be obtained from the EEG. Sterman and Mann (1995) state that EEG frequency changes may be a valid and objective index for mental effort with psychomotor activity, signal processing and intrinsic attentional modulation during complex performance. EEG frequency changes possibly also reveal information about task-related cognitive resource allocation, task mastery and task overload.

Van Orden, Limbert, Makeig, & Jung (2001) found blink frequency, fixation frequency and pupil diameter to be most predictive variables relating eye activity to target density. Moving mean estimation and artificial neural network techniques enable information from multiple eye measures to be combined to produce reliable near-real-time indicators of workload in some visuo-spatial tasks.
Van Orden et al. (2001) state that many eye movement parameters are highly task-specific. They combine multiple eye measures in utilising sufficiently short integration times (about 1 minute), in an attempt to estimate workload real-time. A previous attempt [Van Orden, Jung, & Makeig 2000] was successful in estimating performance and detecting drowsiness in a sustained visual tracking task. Van Orden et al. (2001) investigated eye activity during a task in which memory and visual activity demands varied over time. Estimated task workload (target density) based on an artificial neural network having several eye movement parameters as input correlated highly with actual target density (within-session: R=.75; between session: R=.66). Note that the quality of performance is not taken into account.

4 Experiment

We measured the mental workload with subjective and physiological measures among operators of the Netherlands Royal Navy. This was part of a larger experiment in which a new concept for information presentation for operators on board frigates was evaluated. The aim of this concept (Basic-T) is to improve the situation awareness of the operator and to reduce the number of operators on board navy frigates. In this experiment, experienced operators had to perform complex tasks in three different scenario's (see [vanDelft & Arciszewski, 2004[).

The aim of the workload measure was to get objective information about differences in workload between the scenarios. Furthermore, we explored the possibility to measure state changes during each scenario by comparing the physiological responses with changes in task load during the scenario.

Six Principal Warfare Officers (PWO; in Dutch CCOs) and six Air Defence Operators (ADO; in Dutch LVOs) participated in the experiment. The data of the first CCO and LVO were not used for further analysis due to several problems with the scenario. The data of one LVO could not be analysed due to bad electrodes. Therefore, the analysis was conducted for 5 CCOs and 4 LVOs.

The operators had to perform three scenarios. The first scenario was a standard Anti-Air-Warfare (AAW) for LVO and an Anti-Surface-Warfare (AsuW) for the CCO. The second scenario was a test scenario in which the complexity of the scenario was increased. The third scenario was a combined AAW an AsuW scenario that was the same for both the CCO and LVO.

We measured the heart rate, heart rate variability, respiration frequency, and eye blinks during task performance. These measures are describes in more detail in Appendix B. Table 1 presents an overview of the directions of change of each physiological measure. These measures have been validated in other experiments (e.g. [Veltman & Gaillard, 1998; Veltman & Gaillard, 1996]).

Table 1   Physiological parameters and the direction of change due to high mental workload.

Parameter	Direction of change from low to high workload
Heart rate	↑
Heart rate variability (mid-band)	↓
Heart rate variability (high-band)	↓
Respiratory frequency	↑
Respiratory amplitude	↑
Eye blink frequency	↓
Eye blink duration	↓

Mental workload was also assessed with subjective measures during the experiment. The participants were asked to rate the time pressure on a scale from 1 to 5. They were triggered to give a rating by a tactile device on the wrist that gave a signal with 90-second intervals. The procedure was similar to the workload watch [Boer, 1994]. This score will be referred to as 'workload watch' scores.

The scenarios were taped on video that was replayed directly after the scenario. The participants evaluated the scenario systematically by means of a special purpose software tool that was used before to evaluate the workload among Navy helicopter crews [Veltman et al., 1999].

The participants were asked to indicate when main activities started and ended. These activities were: Situation Awareness (SA), Threat Assessment (TA), Decision Making (DM) and Direction and Control (DC). Furthermore, each minute three rating scales appeared at the computer screen. The participants had to estimate their 'mental effort', 'level of routine actions' and 'time pressure' during the last minute of the scenario. Immediately after the video replay, the scores of the participant was presented as a graph (time line) on the computer screen. Subsequently, the participants were asked to give additional information for segments in which they gave high workload ratings. All data from the video replay sessions and scores from the workload watch are presented in Appendix A.

4.1 Results

The physiological data were analysed in two levels. At the first level, differences between scenarios were analysed and at the second level changes within each scenario were analysed.

4.1.1 Difference between scenarios

Averages of the physiological data for each scenario are calculated for the CCOs and LVOs. Because we did not have baseline values, no conclusions can be drawn about differences in physiological reactions between the CCOs and LVOs. We will only look at differences between the scenarios.

Cardiovascular measures

The results of the cardiovascular measures are presented in Figure 3. For both the CCOs and the LVOs significant differences between the scenarios were found. Post-hoc analysis revealed that the heart rate during scenario 2 was significantly lower than the heart rate in scenario 3. This effect was found for both the CCOs and LVOs. HRV (both the mid- and high-band) was only different for the CCOs. Post-hoc analysis revealed that this was due to the difference between scenario 1 and 3. The HRV in scenario 3 was higher than the HRV in scenario 1.

The results of the HR and HRV are contradictory. The HR results indicate that scenario 3 is the most effortful, while this is not supported by the HRV data. The HRV results indicate that the CCOs invested less effort in scenario 3.
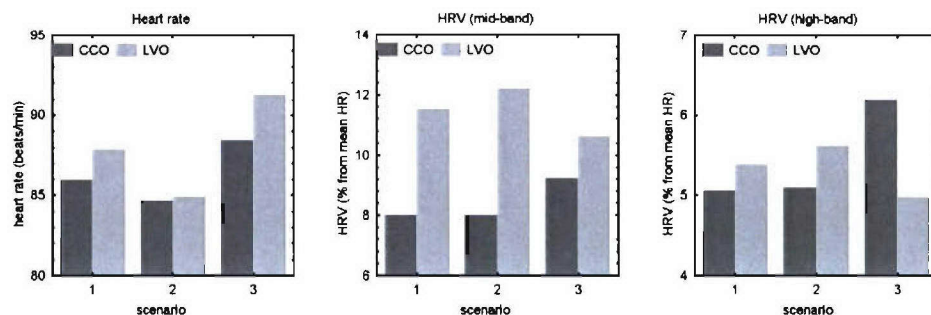

Figure 3   Average cardiovascular activity for the CCO and LVO.

Respiratory measures

The results of the respiratory frequency and amplitude are presented in Figure 4. Statistical analysis revealed no differences between the scenarios.
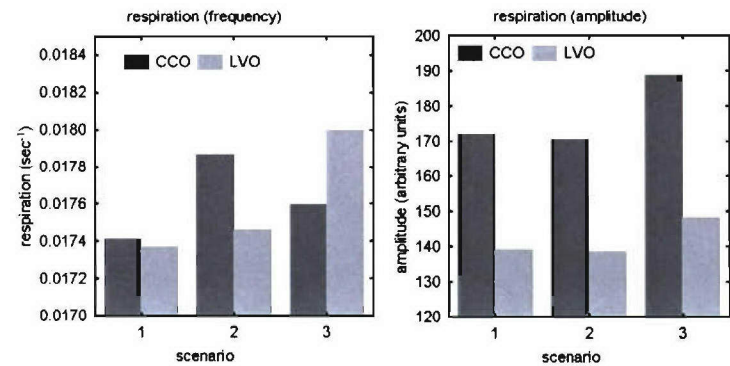

Figure 4   Average respiratory activity for the CCO and LVO.

Eye blinks

The results of the eye blink measurement are presented in Figure 5. Blink amplitude showed significant differences between the scenarios for both the CCOs and LVOs. Post-hoc analysis revealed that this was due to scenario 1 during which the blink amplitude was much higher than during the other scenarios.

Blink amplitude may strongly depend upon the impedance between electrodes that can change during the experiment. Impedance differences instead of workload differences might be the reason for the present statistical effect. Therefore, blink amplitude was not included in the combined workload index (see next section).
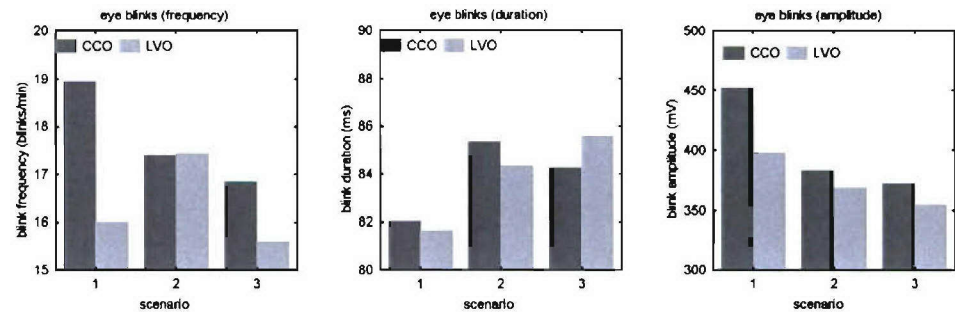
Figure 5   Average blink activity for the CCO and LVO.

4.1.2   *Physiological reactions within the scenarios*

We explored the possibility to calculate a workload indicator based on the physiological parameters. In order to take together the different physiological parameters, we calculated time lines of each parameter. The time resolution is different for each measure. Heart rate frequency, for example, can change within a few seconds, while heart rate variability in the mid-band takes about 15 seconds to change. Blink frequency can change very fast during time segments in which blinks are made but does not change when no blinks are made. We calculated the physiological parameters within the lowest time resolution for each measure and resampled the results with a sampling frequency of 0.25 Hz. (4 seconds). The means that we have a value of each parameter

for each 4-second period. The fast changing parameters might change strongly between two succeeding segments, while the slow changing parameters will not change much. In the next step we assigned workload labels for each parameter and each 4-second segment. These labels were 'normal workload' (value 0), 'moderate workload' (value 1) and 'high workload' (value 2). This step was rather arbitrary, because we did not have absolute criteria for the levels of workload.

We calculated frequency distributions of each measure for all 4-second segments for the three scenarios together. The values that belonged to the 80th percent of the highest values were labelled 'moderate workload' and those values that belonged to the 90th percent of the highest values were labelled 'high workload'. We used the 20th and 10th percentiles for those measures of which low values indicate a high workload.

The next step was to calculate an overall workload estimate for each 4-second segment by adding up the values of each parameter. Because we had seven physiological parameters, the highest workload estimate could be 14 if all parameters pointed towards high workload. The lowest value is 0 when all parameters were labelled 'low workload'.

By applying this procedure we assumed that 80% of the time the workload was normal, 10% of the time the workload of the scenario was moderate and 10% of the total time the workload was high.

Figure 6 presents an example of the combined workload scores of one LVO during scenario 1. The upper graph shows the sum of the workload estimator (sum of the seven parameters) for each 4-second segment. This value can range between 0 (when no parameter indicates a high workload) and 14 (when all parameters indicate a high workload). The lower graph presents the subjective ratings that were assessed during the scenario (workload watch) and the three ratings that were assessed after the scenario (effort, routine and time pressure).

The figure shows that the values of the physiological workload estimates fluctuate considerably. There are several segments with a high overall workload value. However, most segments last only for a short period of time. Between 780 and 820 the overall workload value remained high for a longer period. During this segment the LVO was busy with an attack which normally requires a lot of attention. The subjective ratings also indicate a high workload during this segment. So, there is some evidence that the combined physiological workload value in this segment is related to high workload. We made these plots for all participants and scenarios and found that in most cases the state indicator fluctuated considerably and there was no clear relation between the state indicator and the subjective ratings.

Figure 6 shows another important finding in this study. During the second attack the participant did not give a workload rating (no workload-watch score). Obviously he was too busy performing the task. From a total of 30 scenarios, missing workload watch scores occurred in 11 scenarios (see Appendix A). Seven of these missed ratings were during a period of high information load (such as an attack). This indicates that subjective ratings during a mission are not reliable, especially during time periods in which information about the workload is most important.
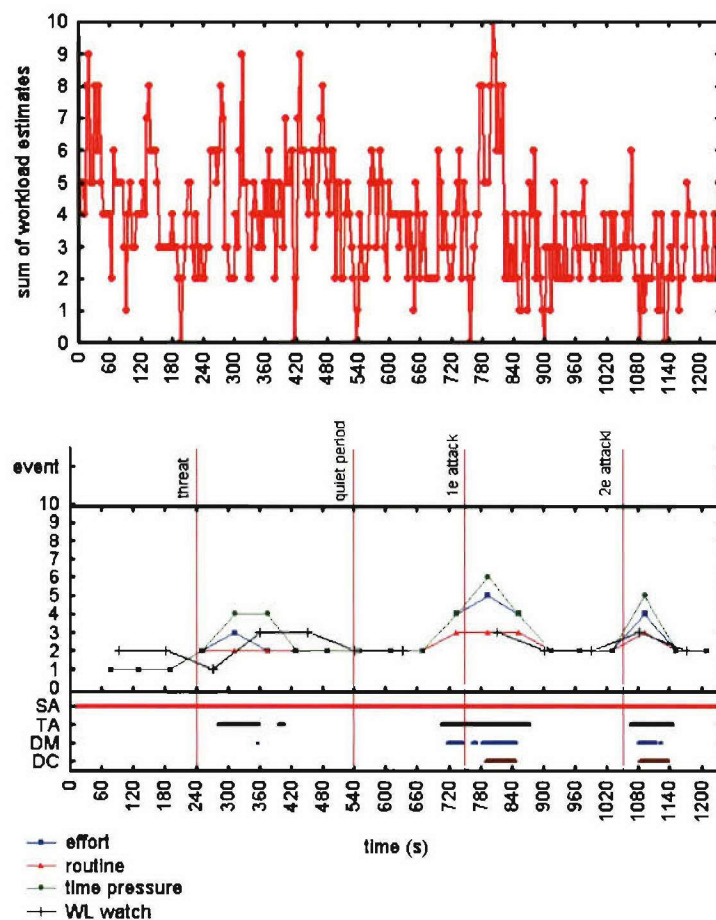
Figure 6   Example of the workload scores of one LVO during scenario 1. The upper graph shows the results of the combined physiological measures in four-second periods. The lower graph shows the subjective ratings during the task (WL watch) and during the video replay after the scenario (effort, routine and time pressure). Furthermore, this graph shows important events during the scenario (vertical lines) and horizontal lines at the bottom of the graph.

We made averages of the state indicator within each scenario to further explore the workload differences of the scenarios. This is similar to what we did with the individual parameters in the previous section, but here the parameters are taken together. Figure 7 shows the averages within each scenario for the LVOs and CCOs. A high value in this figure means that relatively many 4-second values got a moderate or high workload label. Because this is a relative measure, no comparisons can be made of the differences in workload of the CCOs and LVOs.

The combined workload value shows a significant difference between the scenarios of the LVOs. Scenario 3 showed the highest overall workload and scenario 2 the lowest workload. All four LVOs showed the same pattern of results. The combined workload scores of the CCOs did not show a difference between the scenarios.
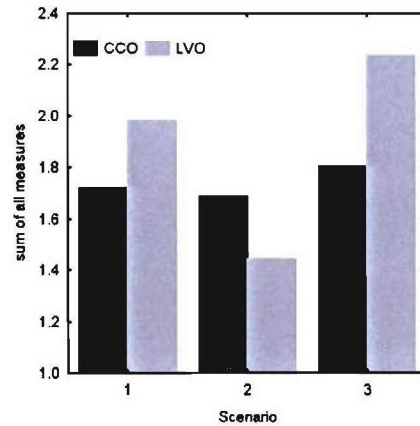
Figure 7   Average workload estimate for each scenario, based on the on the seven physiological parameters.

4.1.3   *Relation between state indicator and number of events*

The number of events in each scenario is determined within 90-second intervals (see [vanDelft et al., 2004]). The number of events provide an indication of the information load. More events lead to a higher information load and it is reasonable to assume that this increases the workload. In order to explore the relation between the events and workload, the state indicator values are averaged within the same 90-seconds periods. The results are presented in Figure 8 and Figure 9. The blue dotted lines present the number of events in each scenario and the five straight lines present the state indicator of each participant. When the state change was a reaction to the number of events we would expect to see the same pattern of results for the number of events and the individual state indicator. However, the figures show that there is no clear relation between the number of events and the overall workload value.
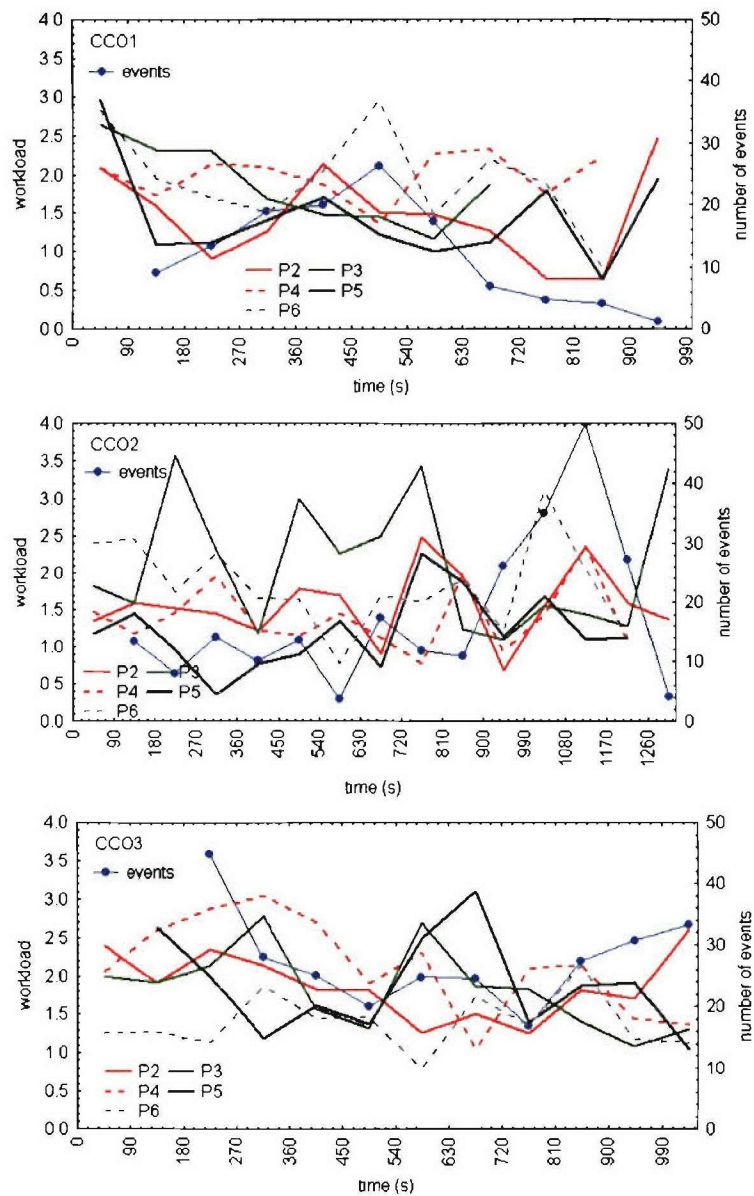
Figure 8   Workload estimates of the CCOs and number of evens during scenario 1, 2 and 3.
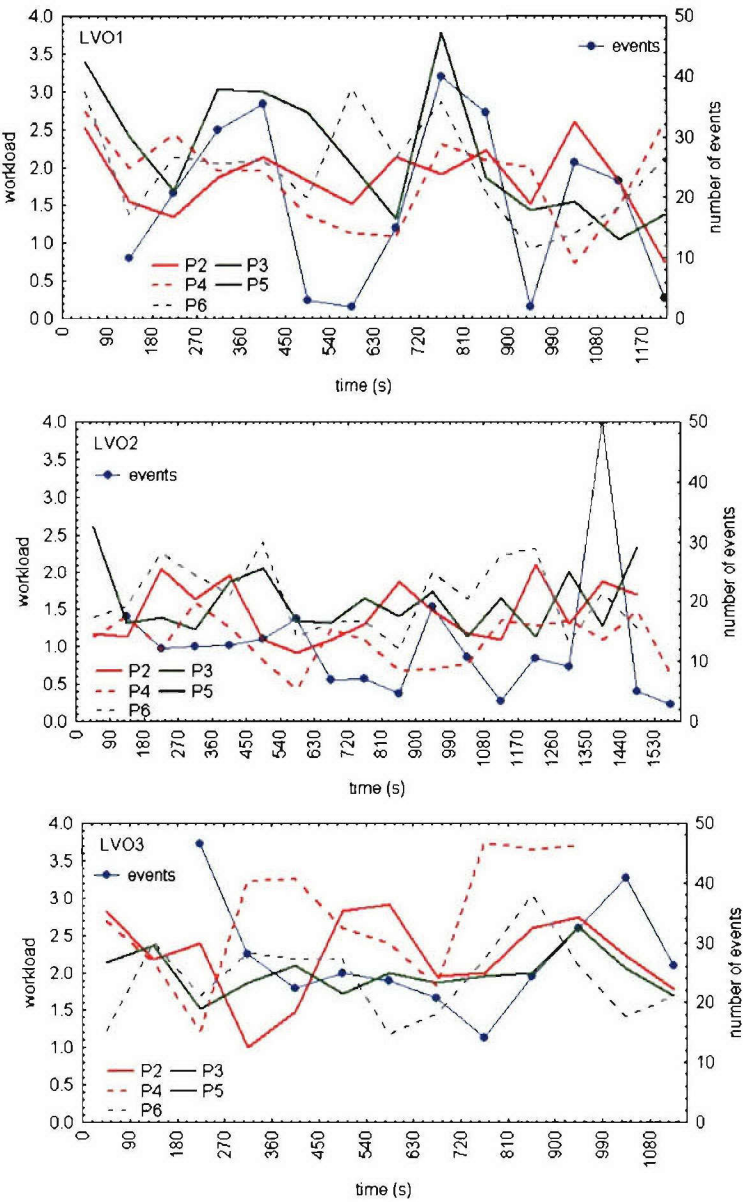
Figure 9   Workload estimates of the LVOs and number of evens during scenario 1, 2 and 3.

5 Discussion

This report provides an overview of adaptive automation and the role of operator state and describes an experiment in which several physiological parameters were measured during complex control tasks. Operator state can be measured with physiological measures such as heart rate, heart rate variability and eye blinks.

The literature review at the beginning of this report shows some promising examples of physiological measures that can be used to adapt the level of automation in order to reduce the workload of the operator. The idea behind using state indictors for adaptive information is rather obvious. If the workload of an operator becomes too high, he/she is more likely to make errors. Decreasing the workload by reducing the task demand will positively affect safety and/or the overall performance. The task load can be reduced when the system takes over some tasks from the operator or when other operators take over some tasks. This requires a re-allocation of tasks that is performed by a dedicated computer system that uses the state information of the operators.

In order to make this work, information about the momentary workload or state of the operator is required. There are several potential problems with such a system. One important aspect is the reliability of the physiological measures. The reliability of the workload assessment must be very high in order not to make wrong decisions. An operator is not likely to accept that tasks are taken away or that he gets more tasks based on wrong assessment of the workload. Such adaptive automation is only likely to work if the operator trusts the system, otherwise he has to check the system which will increase the workload instead of the intended workload reduction.

This report describes an experiment in which the workload is measured with several subjective and physiological measures. Operators had to perform three different scenarios. The workload differences of the scenarios were explored and the changes in workload within the scenarios were investigated.

The physiological data did not show a clear difference between the scenarios. Heart rate indicated that the effort investment of the CCOs and LVOs was relatively low during scenario 2 and relatively high during scenario 3. Heart rate variability (HRV) however indicated that LVOs invested less mental effort in scenario 3. It should be noted that interpreting HRV in a condition in which the operator speaks a lot is rather difficult. Normally, the HRV will decrease when the operator invests more effort into a task. However, changing the respiratory pattern does affect the HRV strongly. Speaking is characterised by a short inhalation followed by a long expiration. This increased the HRV substantially and can mask the effects of the higher effort investment.

This makes HRV less applicable for operator tasks, unless speaking is taken into account (e.g., use the data in segments in which the operator is not speaking). This must be taken into account in future experiments.

A state indicator was calculated that was based on all physiological parameters. This indicator showed no differences between the scenarios for the CCOs. For the LVOs the state indicator showed a relative low effort investment in scenario 2 and a relative high effort investment in scenario 3. All participants showed the same pattern of results. This makes it reasonable to assume that a combination of measures is more robust than a single measure. Scenario 1 was intended to be the least demanding. However, the physiological results show that the LVOs invested the lowest amount of effort during scenario 2. The following two arguments can be used to explain the relative low effort investment of the LVOs in scenario 2:

1 it appeared that the number of tactical events was lower in scenario 2 compared to scenario 1 (see [vanDelft et al., 2004]);

2 it is known that participants are more aroused during a first scenario, because they do not know exactly what to expect.

It must be noted that other indicators show that differences in workload between the scenarios was low. The workload was in the normal range for all three scenarios (see [vanDelft et al., 2004]) and therefore, large differences in physiological reactions are not likely to occur.

The physiological reactions within the scenarios did not show clear results. The combination of physiological measures showed a large variation that did not show a clear relation with other workload indicators such as the subjective ratings during the experiment (workload watch) and the subjective ratings after the experiment. For this analysis it should also be noted that the task load differences in the scenario were small. The participants were continuously busy performing the tasks and it did not occur that the task became too difficult. Even when there were not many events, the operators had to spend all their attention to the task in order to anticipate to new events.

Evaluating physiological workload measures is rather difficult because there is often no absolute reference value. Most often, the task demand is used as a reference. When the task demand increases, the effort investment is expected to increase too. Such setup was also used in the present experiment. If there is a relation between changes in task demand and the workload measure in the expected direction, there is strong evidence that the workload measure is valid. However, when the experiment does not show such a relation there is no direct reason to assume that the workload measure is not valid.
A factor that complicates the relation between task load and physiological responses is the complexity of a task. Operators have to build up a mental model of the environment and use this model to evaluate new information. As long as new information is congruent to the mental model, the operator does not have to pay much attention to this information. However, when the information does not fit to the mental model, the operator has to perform many additional checks in order to adapt the model or to decide that the information is not valid [Veltman et al., 1999]. This can happen independent of the objective task load (number and difficulty of tasks to be performed). Even if the task load is low, there may be some information that does not fit to the mental model and causes a high workload.
In complex tasks it is more likely that the mental model does not completely match to the environment and that additional checks have to be performed. Thus, the fact that there was no clear relation between the task load and the physiological responses does not necessarily mean that the physiological measures are not valid to be used as an online state indicator.

Physiological measures provide continuous information. In the present experiment all parameters were calculated in four-second periods. The results showed that variation between successive periods was rather high. It is not reasonable to assume that the mental workload varied vast in this experiment. It was argued earlier that it is likely that the effort expenditure was rather constant in this experiment. This means that it is not reasonable to calculate the state estimator within such small time segments. It is difficult to extract an optimal time window for averaging the physiological measures from the present data.

The present experiment does not provide a clear answer to the question whether physiological measures can be used to estimate the state of an operator. There are some other important questions to be answered before physiological measures can be used in adaptive automation. For example: in what situations should the workload of an operator be reduced if it is too high? This question is crucial for the concept of adaptive automation. The theoretical model that is described in this report shows that there is a continuous interaction between an operator and the (task) environment. The operator normally is trying to adapt to the requirements of the task by regulating the effort expenditure or by changing the task goals. If the state of the operator can be measured continuously and this is used to adapt the task demands, than there is a situation in which there are two adaptive systems working together: the operator trying to adapt to the task environment and a task environment that is trying to adapt to the state of the operator. Although such a system is not tested yet, it is not likely that such a system will work properly. As long as the operator is able to adapt to the changing task demands he should not be disturbed by an adaptive system that is changing his main task.

Thus, in normal situations, a system should not change the main task when the operator invests a lot of mental effort. There are, however, several circumstances in which state estimators might be valuable in adaptive automation. For example, a system might delay information that is not very important and not time critical when the operator is highly loaded. The system might also present information of new tasks to another operator.

Another example in which adaptive automation might work is a situation in which the operator does not show adaptive behaviour any more. In this situation it might be better to take over control. Examples of such situations are when the task becomes very demanding and the operator does not increase the mental effort investment or when the operator appears to invest a lot of mental effort during tasks that are not very demanding. More information is required for such a system than information about the state of the operator. It is also necessary to have information about the context (e.g., how important is the task) and the task demands. There is not much literature about indicators of non-adaptive behaviour. EEG parameters might provide information about such a state. Several studies in which EEG parameters are used for adaptive automation (e.g. [Scerbo, Freeman, & Mikulka, 2000]) show very promising results in adaptive automation. These authors use frequency components to allocate tasks to the operator or computer (they use the amplitude of high frequency EEG divided by the amplitude of low-frequency EEG). It appears that such a system works if operators get more tasks when they show relative little high-frequency EEG activity. Low-frequency EEG is related to a state of low alertness, which can be interpreted as a non-adaptive state. Giving more tasks in such a state results in a higher involvement of the operator. These studies are mainly conducted in low-workload situations.

As pointed out above, it is not likely that physiological measures can be used in adaptive automation within small time segments (real time). However, psychophysiological parameters might be useful when the state is calculated in much longer time intervals (near real time). When the operator is investing a lot of effort for a substantial time period, he will become fatigued and as a consequence he is likely to make errors. It is positive for the long-term task performance and for the well being of the operator if this can be detected and operators get the possibility to reduce the workload to recover from the high-workload period.

Objective information about the workload within larger time segments is also useful when new systems or new concepts of task distribution have to be evaluated. It is not

only important to look at the performance but also important to look at the costs such as the effort investment.

A disadvantage of physiological measures is that sensors have to be attached to the operator. New techniques make it possible to get objective information about the state without attaching electrodes. A temperature change of the face that can be measured with an infrared camera that is positioned in front of the operator is an example of such a technique [Vos and Veltman, 2005]. Information about eye activity such as eye blinks can also be obtained without attaching sensors. The newest eye tracking systems use cameras that can be positioned below the computer screens of an operator. These systems not only provide information about the eye point of gaze but also about the blink frequency and the diameter of the pupil. Sensors that have to be attached to the operator also become smaller and wireless communication becomes standard for many sensors such as heart rate sensors. This makes physiological sensors more applicable in applied settings.

There were some other relevant findings in the present experiment. The data showed that subjective ratings during the task performance does not always provide direct information about mental workload. It happened very often that the operator did not give a workload rating during segments of a high task load. Omissions of subjective ratings may provide an indication of high workload. Physiological measures do not have this problem. They continuously provide information about the state of the operator.

6   Conclusions

The literature review shows that physiological parameters that can measure the state of an operator are very promising to be used for adaptive automation. From a theoretical point of view and from experimental data there are good reasons to be less optimistic because:

- An operator is continuously adapting to changes in task load. Physiological reactions are a sign of this adaptive behaviour. If a system uses this information to reduce the taskload, there are two adapting systems that can work counterproductive.
- The state of an operator must be measured very reliably in order to be used in adaptive automation. The present experiment shows that this is hard to achieve.
- Physiological measures might be useful in situations in which the operator is not adaptive anymore (e.g. when he is overloaded or when he does not react to changes in task demands). This information might be more useful for adaptive information than changes within a normal range of states.

Additional conclusions from the experiment:
- The use of subjective workload measures during complex task performance is questionable because operators tend to skip the ratings when the task becomes difficult.
- Averaging physiological parameters within four-second windows results in noisy data. Longer time windows are needed to get a stable estimation of the state, but this makes it less applicable for adaptive automation.

7   References

Bainbridge, L. (1983). Ironies of automation. *Automatica, 19,* 775-779.

Boer, L. C. (1994). *Workloadwatch* (Rep. No. SIG 1994 B 04).

Byrne, E.A., Chun, K., Hilburn, B., Molloy, R.& Parasuraman, R. (1994). Effect of tracking difficulty on secondary task performance, heart rate variability, and subjective perceptions. Paper presented at the Annual Meeting of the Society for Psychophysiological Research, Atlanta, GA.

Byrne, E. A. & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological Psychology, 42,* 249-268.

Carver, C. S. & Scheier, M. F. (1998). *On the Self-Regulation of Behavior.* Cambridge: University Press.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37,* 32-64.

Endsley, M. R. (1996). Automation and situation awareness. In R. Parasuraman & M. Mouloua (Eds.), *Automation and Human Performance: Theory and Application* (pp. 163-181). Mahwah, NJ: Erlbaum.

Ephrath, A. R. & Young, L. R. (1981). Monitoring vs. man-in-the-loop detection of aircraft control failures. In J. Rasmussen & W. B. Rouse (Eds.), *Human detection and diagnosis of system failures* (pp. 143-154). New York: Plenum Press.

Farwell, L. A. and Donchin, E. (1988) Talking Off The Top Of Your Head: A Mental Prosthesis Utilizing Event-Related Brain Potentials. Electroencephalography and Clinical Neurophysiology, 70: 510-513.

Farrell, S. & Lewandowsky, S. (2000). A connectionist model of complacency and adaptive recovery under automation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26,* 395-410.

Fitts, P. M. (1951). *Human engineering for an effective air navigation and traffic control system.* Washington, DC: National Research Council.

Gehring WJ, Goss B, Coles MG, Meyer DE, Donchin E (1993) A neural system for error detection and compensation. Psychol Sci 4:385–390.

Harris, W. C., Hancock, P. A., Arthur, E., & Caird, J. K. (1991). *Automation influences on performance, workload, and fatigue.* In .

Hendy, K. C., East, K. P., & Farrell, P. S. E. (2001). An information processing model of operator stress and performance. In P.A.Hancock & P. A. Desmond (Eds.), *Stress, workload and fatigue* (pp. 34-82). Mahwah (NJ), London: Lawrence Erlbaum Associates, Publishers.

Hilburn, B., Molloy, R., Wong, D., & Parasuraman, R. (1993). Operator versus computer control of adaptive automation. In . Columbus OH: Department of Aviation, The Ohio State University.(pp. 161-166)

Hoc, J.-M. (2000). From human-machine interaction to human-machine cooperation. *Ergonomics, 43,* 833-843.

Hockey, G. R. J. (2003). Operator Functional State as a Framework for the Assessment of Performance Degradation. In G.R.J.Hockey, A. W. K. Gaillard, & A. Burov (Eds.), *Operator Functional State: The assessment and Prediction of Human Performance Degradation in Complex Tasks.* (pp. 8-23). Amsterdam: IOS Press, NATO Science series.

Hockey, G. R. J., Gaillard, A. W. K., & Burov, O. (2003). *Operator functional state - The assessment and prediction of human performance degradation in complex tasks.* Amsterdam: IOS Press.

Holding, D. H. (1983). Fatigue. In G.R.J.Hockey (Ed.), (Dordrecht: Martinus Nijhoff.

Jain, A., Martens, W. L. J., Mutz, G., & Weiß, R. K. (1996). Towards a Comprehensive Technology for Recording and Analysis of Multiple Physiological Parameters within their Behavioral and Environmental Context. In J.Fahrenberg & M. Myrtek (Eds.), *Ambulatory Assessment. Computer-assisted Psychological and Psychohysiological methods in monitoring and field studies* (pp. 215-235). Seatle Hogrefe: Huber, Publisher.

Jorna, P.G.A.M. (1992). Spectral analysis of heart rate and psychological state: a review of its validity as a workload index. Biological Psychology, 34, 237-257.

Kaber, D. B. & Riley, J. M. (1999). Adaptive automation of a dynamic control task based on secondary task workload measurement. *International Journal of Cognitive Ergonomics, 3,* 169-187.

Kessel, C. J. & Wickens, C. D. (1982). The transfer of failure-detection skills between monitoring and controlling dynamic systems. *Human Factors, 24,* 49-60.

Kirlik, A. (1993). Modeling strategic behavior in human-automation interaction: Why an 'aid' can (and should) go unused. *Human Factors, 35,* 221-242.

Kramer, A. F., Trejo, L. J., & Humphrey, D. (1996). Psychophysiological measures of human information processing activities: Potential applications to automated systems. In R. Parasuraman & J. Mouloua (Eds.), Automation and Human Performance: Theory and Applications. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 137-162.

Martens, W. L. J. (1992). The fast time frequency transform (F.T.F.T): a novel online approach to the instantaneous spectrum. In *Proceedings of 14th International Conference of the IEEE Engineering in Medicine and Biology Society.* Paris:

Morrison, J. & Gluckman, J. P. (1994). Definitions and prospective guidelines for the application of adaptive automation. In M. Mouloua & R. Parasuraman (Eds.), *Human performance in automated systems: Current research and trends* (pp. 256-263). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mouloua, M., Parasuraman, R., & Molloy, R. (1993). Monitoring automation failures: Effects of single and multi-adaptive function allocation. In .(pp. 1-5)

Papanicolaou, A.C., Johnstone, J. Probe evoked potentials: Theory, method and applications. Intl. J. Neurosci. 24:107-131, 1984.

Parasuraman, R., Bahri, T., Deaton, J., Morrison, J., & Barnes, M. (1992). *Theory and design of adaptive automation in aviation systems* (Rep. No. NAWCADWAR-92033-60). Warminster, PA: Naval Air Warfare Center.

Parasuraman, R., Molloy, R., Mouloua, M., & Hilburn, B. (1996). Monitoring of automated systems. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: theory and application* (pp. 91-115). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Parasuraman, R., Molloy, R., & Singh, I. (1993). Performance consequences of automation-induced 'complacency. *International Journal of Aviation Psychology, 3*, 1-23.

Parasuraman, R., Mouloua, M., & Hilburn, B. (1999). Adaptive aiding and adaptive task allocation enhance human-machine interaction. In M. Scerbo & M. Mouloua (Eds.), *Automation Technology and Human Performance: current research and trends* (pp. 119-123). Mahwah, NJ: Erlbaum.

Parasuraman, R. & Riley, V. A. (1997). Humans and automation: use, misuse, disuse, abuse. *Human Factors, 39*, 230-253.

Parasuraman, R. & Mouloua, M. (1996). *Automation and human performance: Theory and applications*. (vols. 1996). xx, 514pp.).

Parasuraman, R., Mouloua, M., & Molloy, R. (1996). Effects of adaptive task allocation on monitoring of automated systems. *Human Factors, 38*, 665-679.

Prinzel, L. J., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., & Pope, A. T. (1998). An assessment of the efficacy of event-related potentials for adaptive automation. Proceedings of the 42nd Annual Meeting of the Human Factors and Ergonomics Society, 1602.

Powers, W. T. (1973). *Behavior: The Control of Perception*. Chicago: Aldine.

Sarter, N. B. & Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. Human Factors 37[1], 5-19.

Satchell, P. (1993). *Cockpit monitoring and alerting systems*. Aldershote, UK: Ashgate.

Scallen, S., F, Hancock, P., A, Duley, J., & A (1995). Pilot performance and preference for short cycles of automation in adaptive function allocation. *Applied Ergonomics, 26*, 397-403.

Scerbo, M., Freeman, F. G., Mikulka, P. J., Parasuraman, R., DI Nocero, F., & Prinzel, L. J. (2001). *The efficacy of psychophysiological measures for implementing adaptive technology* (Rep. No. L-18086; NASA?TP-2001-211018). Hampton, Virginia: National Aeronautics and Space Administration, Langley Research Center.

Scerbo, M. W., Freeman, F. G., & Mikulka, P. J. (2000). A biocybernetic system for adaptive automation. In R.W. Backs & W. Boucsein (Eds.), *Engineering Psychophysiology: Issues and applications* (pp. 241-254). Mahwah (N.J.), London, Lawrence Erlbaum Associates.

Scerbo, M. W. (1996). Theoretical perspectives on adaptive automation. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications. Human factors in transportation* (pp. 37-63). xx, 514pp.

Sterman & Mann (1995). Biological Psychology, 40 (1-2), 115-130.

vanDelft, J. H. & Arciszewski, H. F. R. (2004). *Eindevaluatie automatiserings- en ondersteuningsconcepten Studie Commandovoering.* (Rep. No. TM-04-A067 | FEL-04-A271). Soesterberg. TNO Human Factors.

Van Orden K., Limbert W., Makeig S., and Jung T-P, Eye Activity Correlates of Workload during a Visualspatial Memory Task, Human Factors, 43(1):111-21, 2001.

Van Orden K., Jung T.P., and Makeig S. (2000). Combined eye activity measures accurately estimate changes in sustained visual task performance. Biological Psychology, 52(3):221-4.

Velden, M. & Graham, F. K. (1988). Depicting heart rate over real time: two procedures that are mathematically identical. *Journal of Psychophysiology, 2,* 291-292.

Veltman, J. A. & Gaillard, A. W. K. (1996). Physiological indices of workload in a simulated flight task. *Biological Psychology, 42,* 323-342.

Veltman, J. A. & Gaillard, A. W. K. (1998). Physiological workload reactions to increasing levels of task difficulty. *Ergonomics, 5,* 656-669.

Veltman, J. A. & Gaillard, A. W. K. (1999). *Mental workload of the tactical co-ordinator of the Lynx helicopter.* (Rep. No. TM-A-036). Soesterberg, The Netherlands: TNO Human Factors.

Veltman, J. A. & Jansen, C. (2003). Differentiation of mental effort measures: consequences for adaptive automation. In G.R.J.Hockey, A. W. K. Gaillard, & O. Burov (Eds.), *Operator Functional State: The Assessment and Prediction of Human Performance Degradation in Complex Tasks.* (pp. 249-259). Amsterdam: IOS Press.

Vos , J.A. & Veltman, J.A. (2005). Mental workload assessment with an infrared camera. (Rep. TNO-DV3 2005-B 003). Soesterberg, The Netherlands: TNO Human Factors.

Whinnery, J. E., Glaister, D. H., & Burton, R. R. (1987). +Gz-induced loss of consciousness and aircraft recovery. *Aviation, Space, and Environmental Medicine, 58,* 600-603.

Wickens, C. D. & Kessel, C. J. (1979). The effects of participatory mode and task load on the detection of dynamic system failures. *IEEE Transactions on Systems, Man & Cybernetics, SMC-9,* 24-34.

Wiener, E. L. & Curry, R. E. (1980). Flight-deck automation: promises and problems. *Ergonomics, 23,* 995-1011.

Wilson, G. F. (1993). Air-to-ground training missions: a psychophysiological workload analysis. *Ergonomics, 36,* 1071-1087.

Wilson, G. F., Fullenkamp, P., & Davis, I. (1994). Evoked potential, cardiac, blink and respiration measures of pilot workload in air-to-ground missions. Aviation, Space, and Environmental Medicine, 65, 100-105.

Wilson, G. F., Lambert, J. D., & Russell, C. A. (2000). Performance enhancement with real-time physiologically controlled adaptive automation.

Yamamoto, Y. & Isshiki, H. (1992). Instrument for controlling drowsiness using galvanic skin reflex. *Medical & Biological Engineering & Computing, 30,* 562-564.

8 Signature

Soesterberg, January 2006
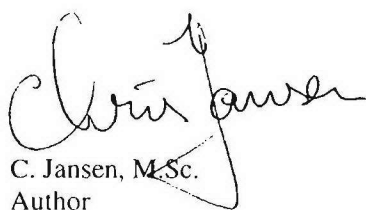
Dr. H.A.H.C. van Veen
Group leader

TNO Defence, Security and Safety

J.A. Veltman, M.Sc.
Project leader/Author

C. Jansen, M.Sc.
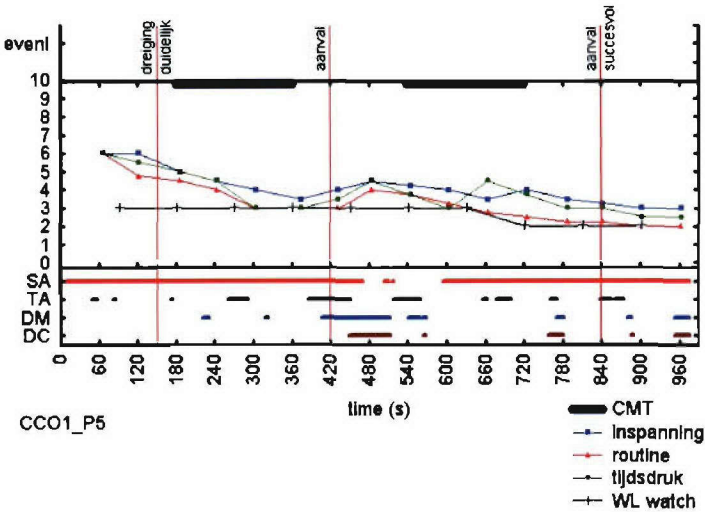Author

A   Time lines of tasks and subjective workload

These figures show the results of the video replay analysis and the workload watch sores that were obtained during the scenarios. The vertical lines indicate important events that were indicated by the participants during and after the video replay.
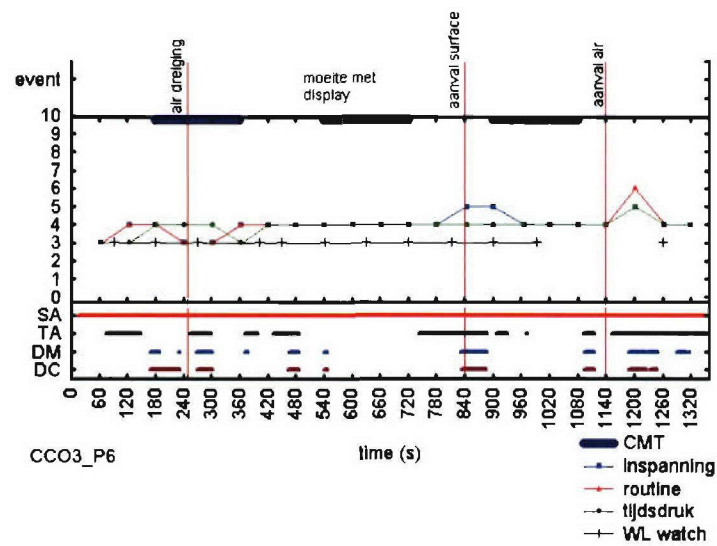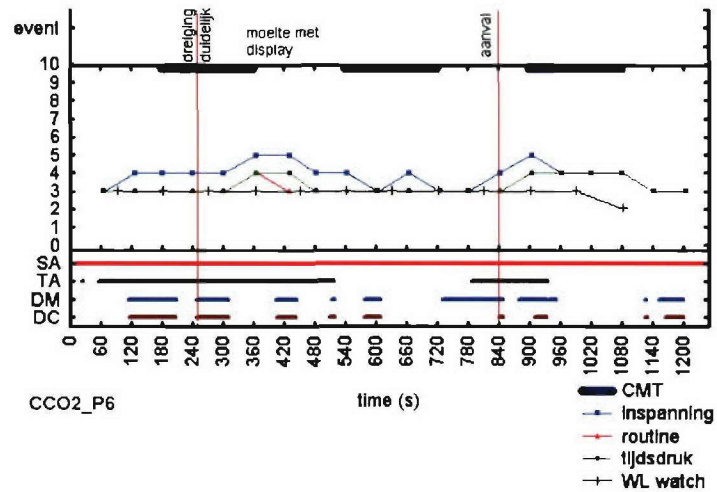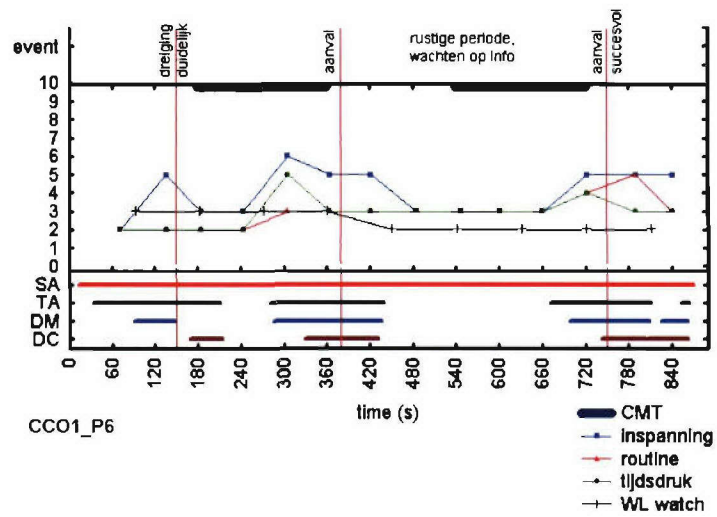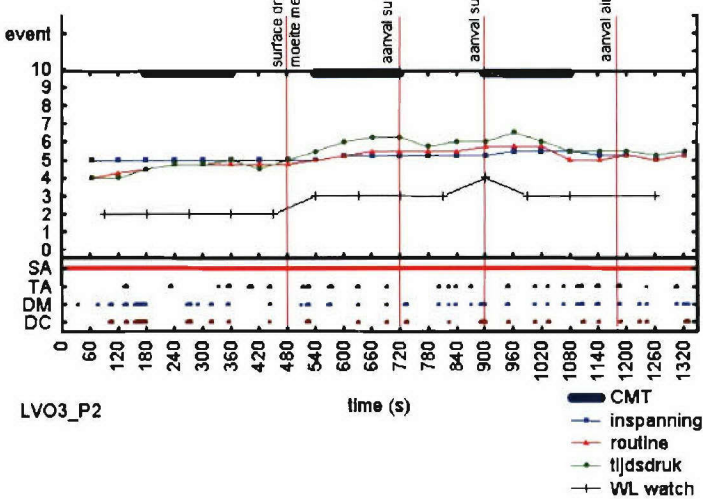

CCO1_P2

CCO2_P2

CCO3_P2

CCO1_P3

CCO2_P3

CCO3_P3

CCO1_P4

CCO2_P4

CCO3_P4

CCO1_P5

CCO2_P5

CCO3_P5

CCO1_P6

CCO2_P6

CCO3_P6

LVO1_P2

CMT	
Inspanning	
routine	
tijdsdruk	
WL watch	

LVO2_P2

CMT	
inspanning	
routine	
tijdsdruk	
WL watch	

LVO3_P2

CMT	
inspanning	
routine	
tijdsdruk	
WL watch	

LVO1_P3

LVO2_P3

LVO3_P3

LVO1_P4

LVO2_P4

LVO3_P4

LVO1_P5

LVO2_P5

LVO3_P5

LVO1_P6

LVO2_P6

LVO3_P6
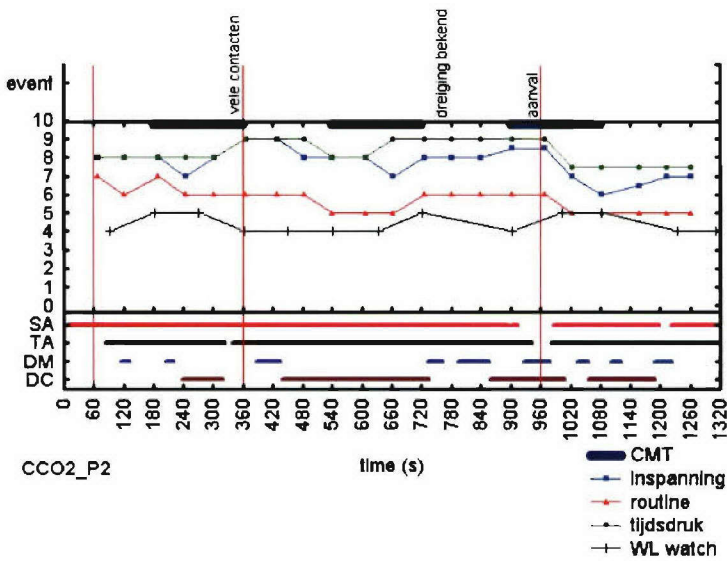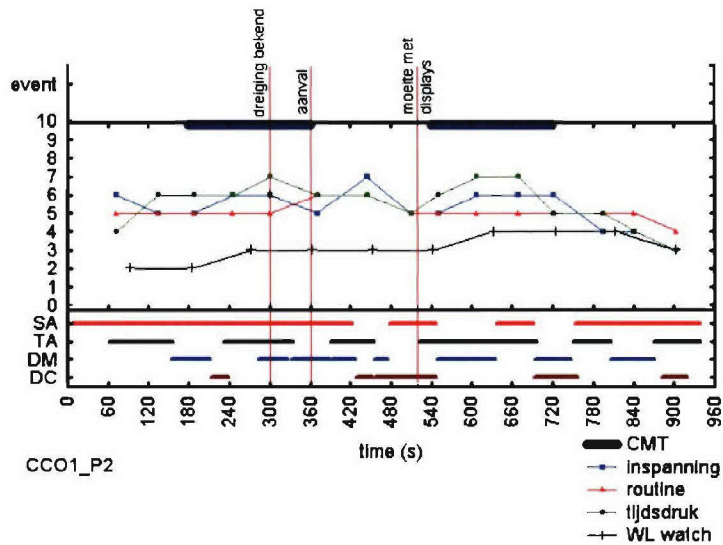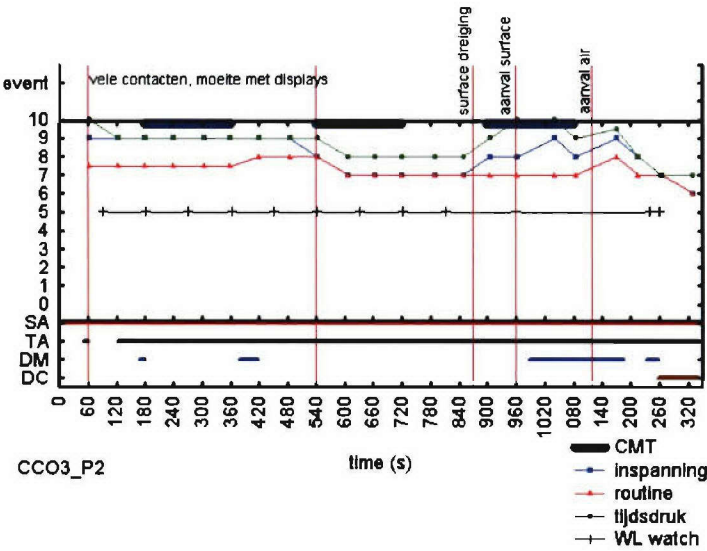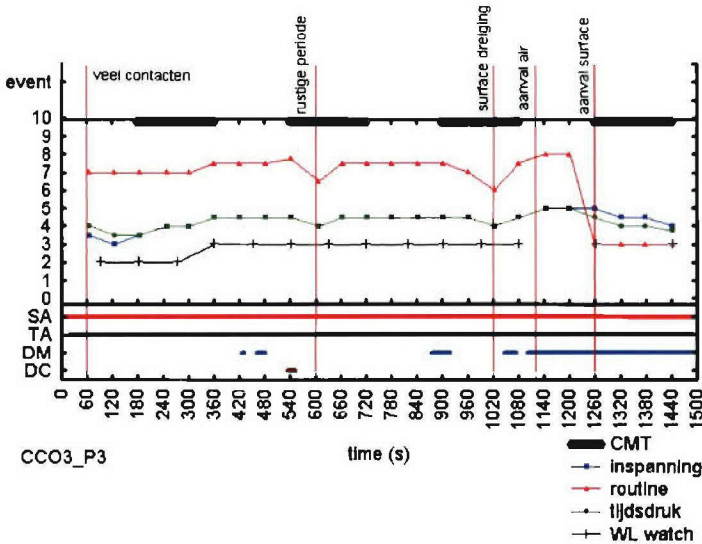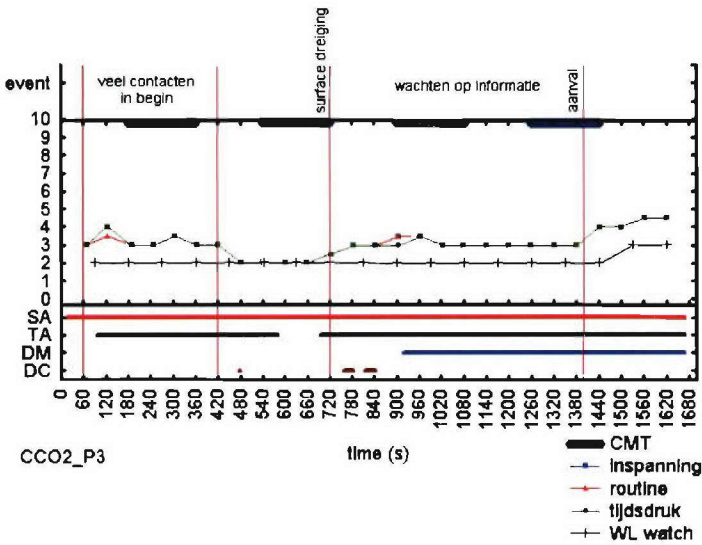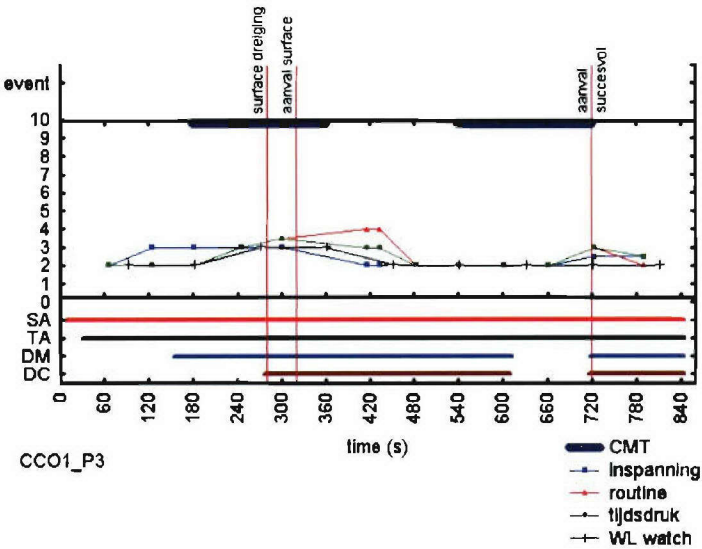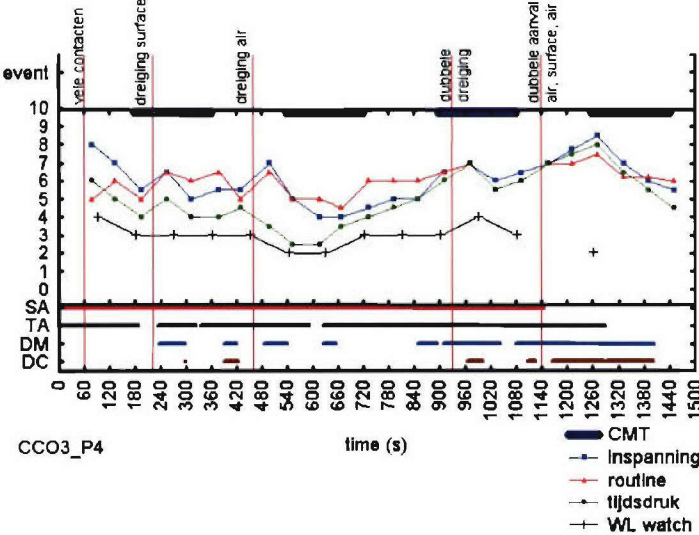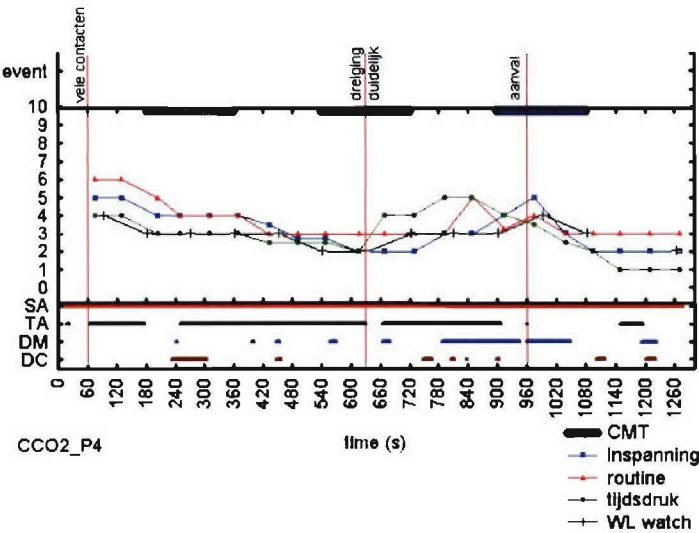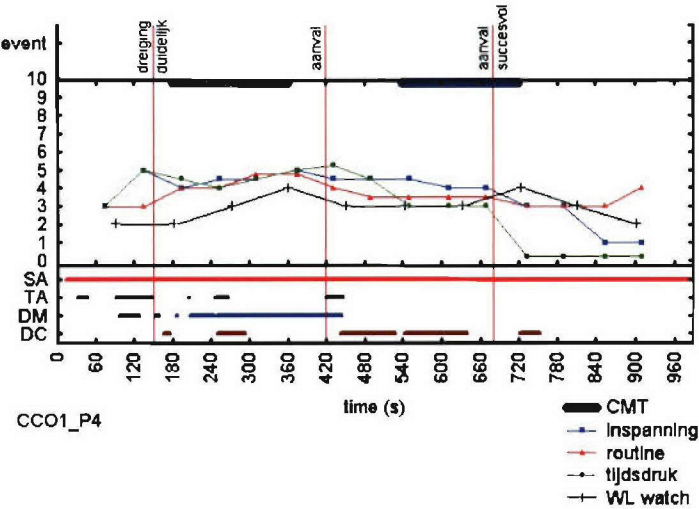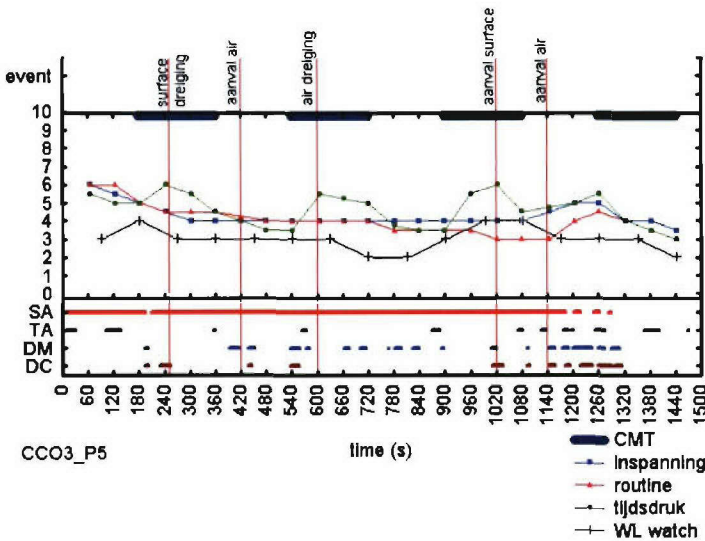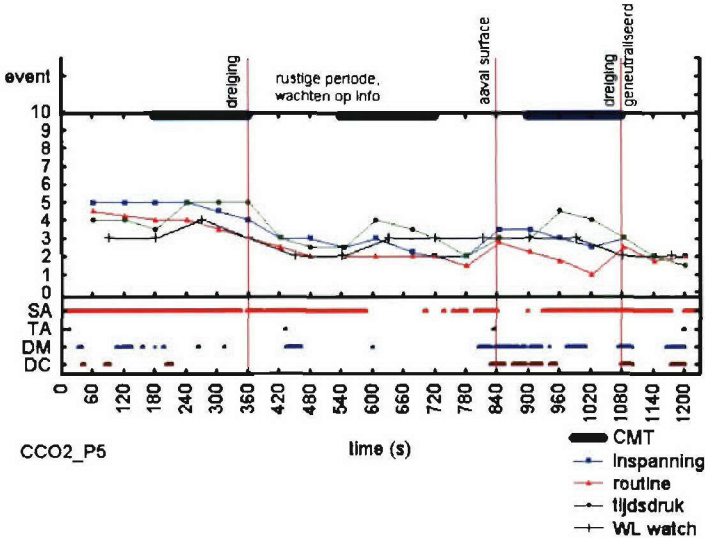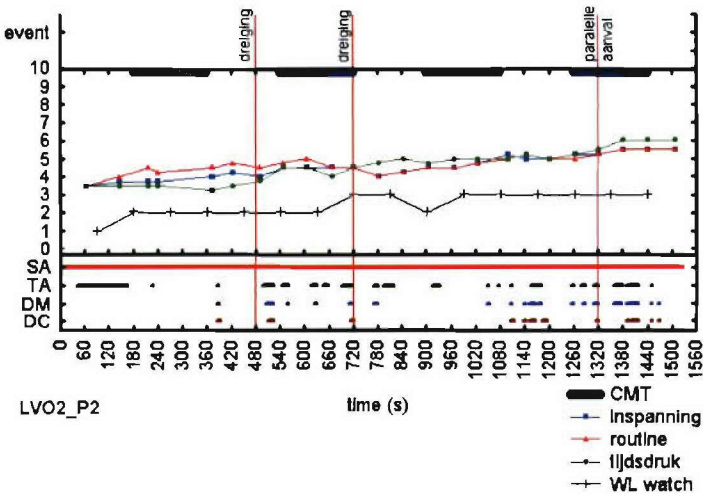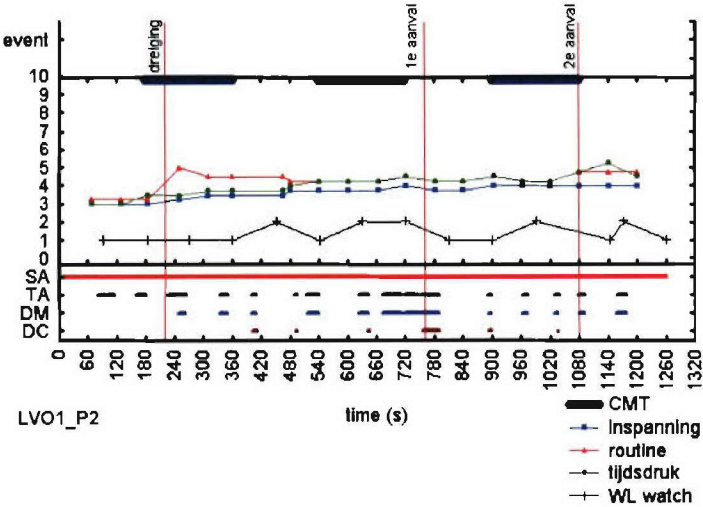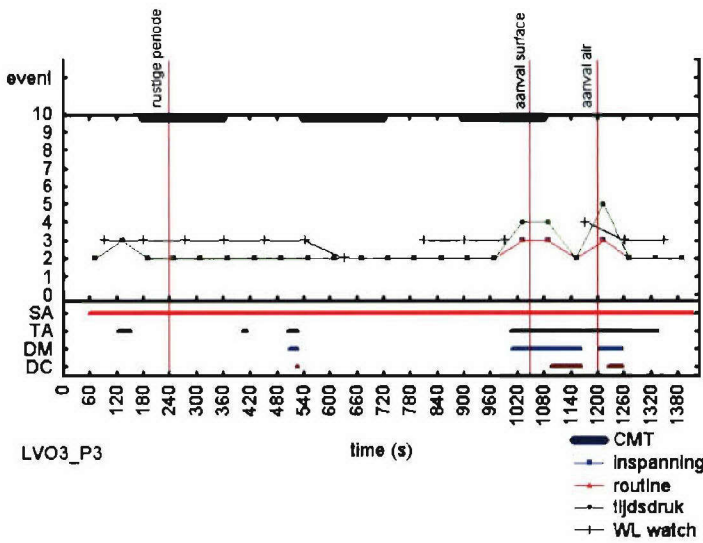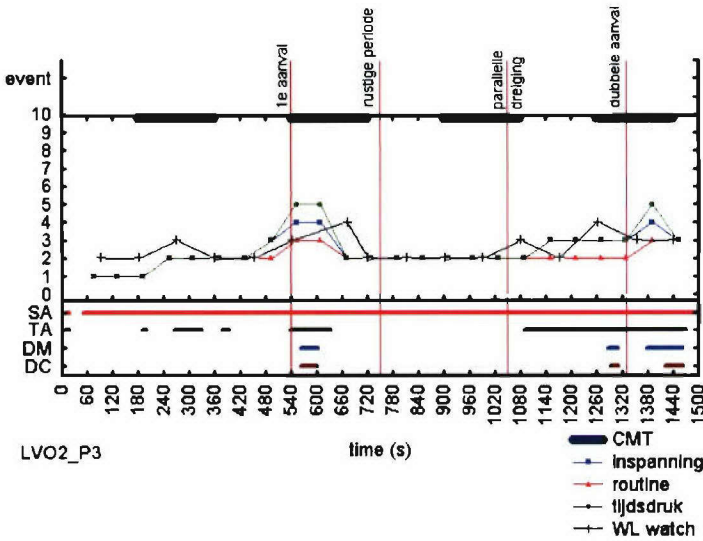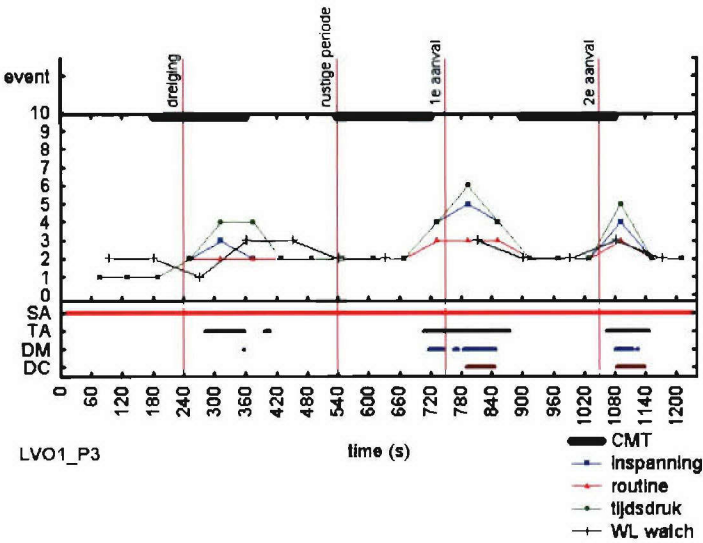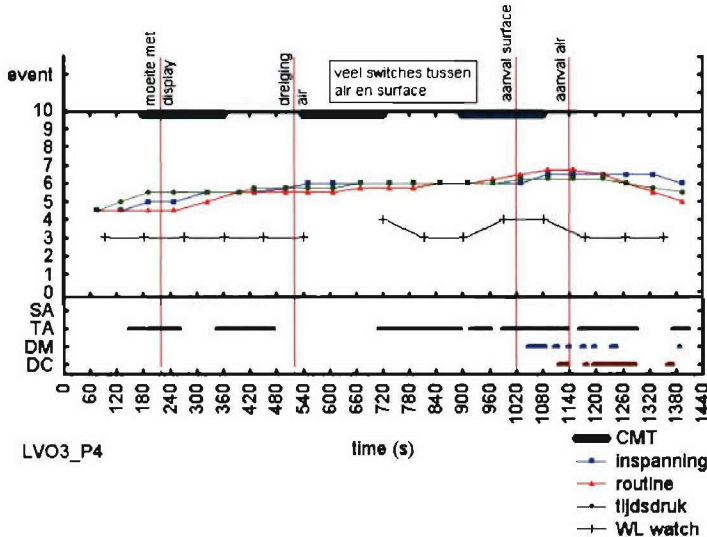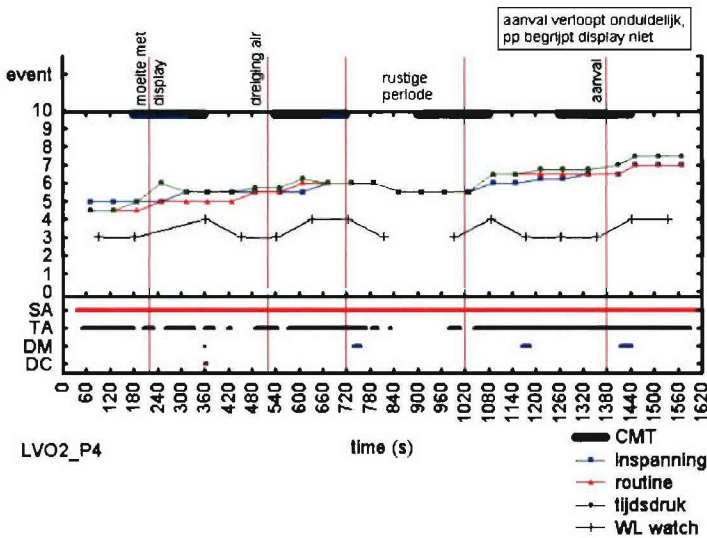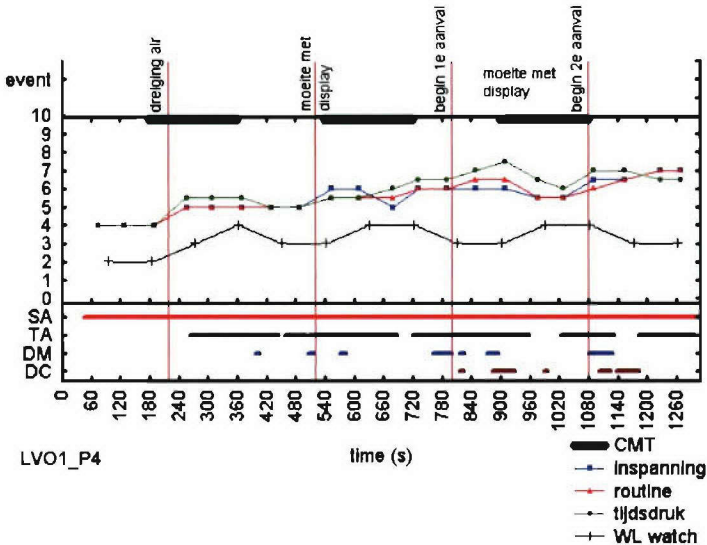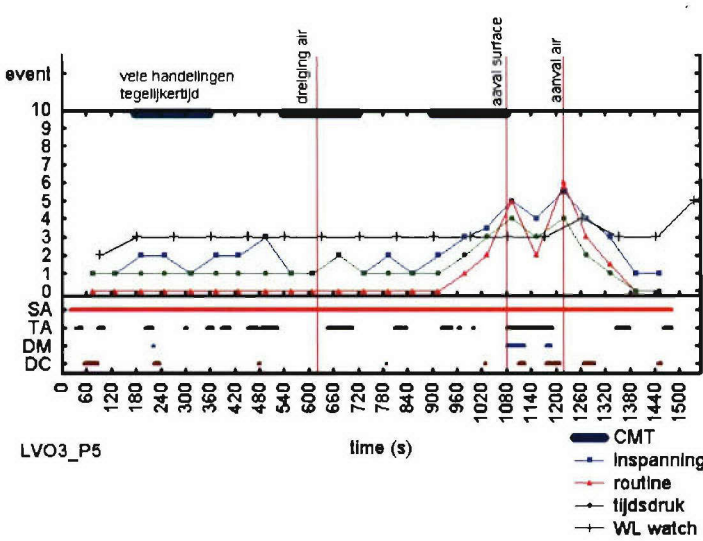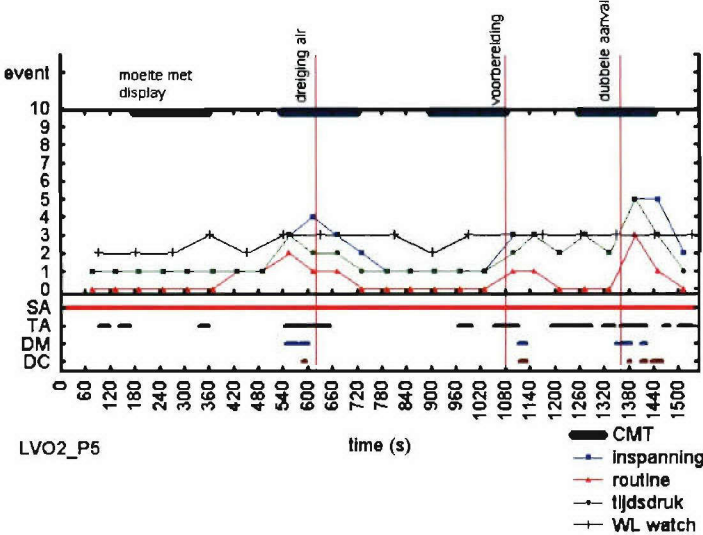
B   Physiological measures

Physiological signals were recorded with a Vitaport II measurement system [Jain et al., 1996].

Several physiological signals were recorded during the experiment. Table 2 presents an overview of the signals that were recorded as well as the measures that were derived from these signals.

Table 2   Physiological signals.

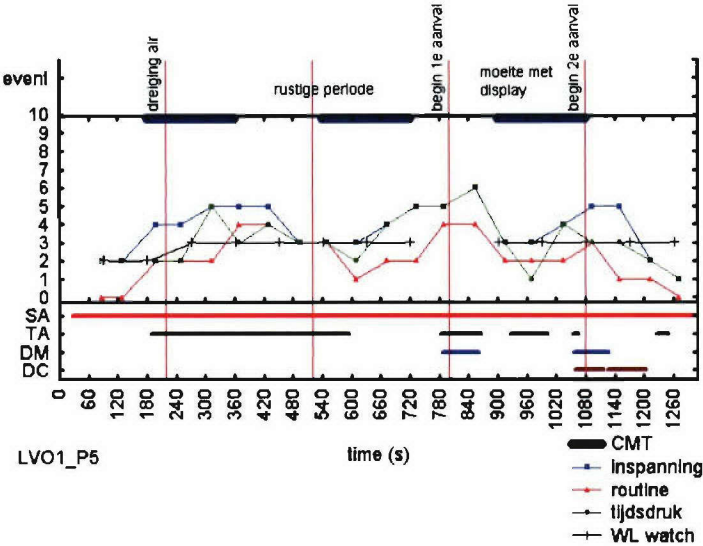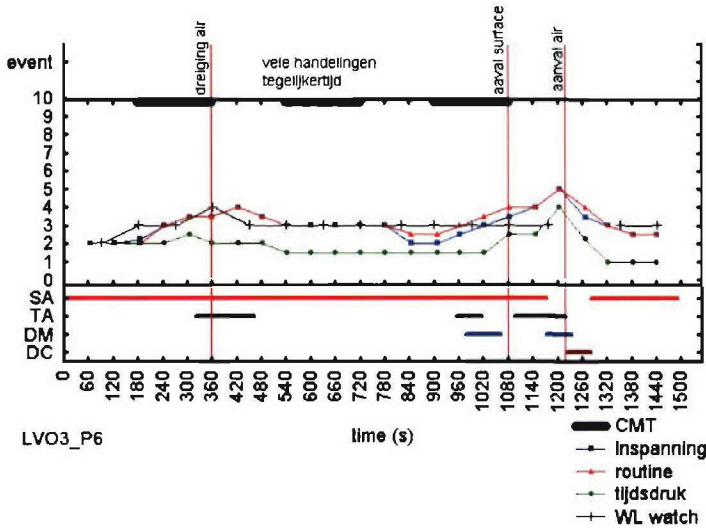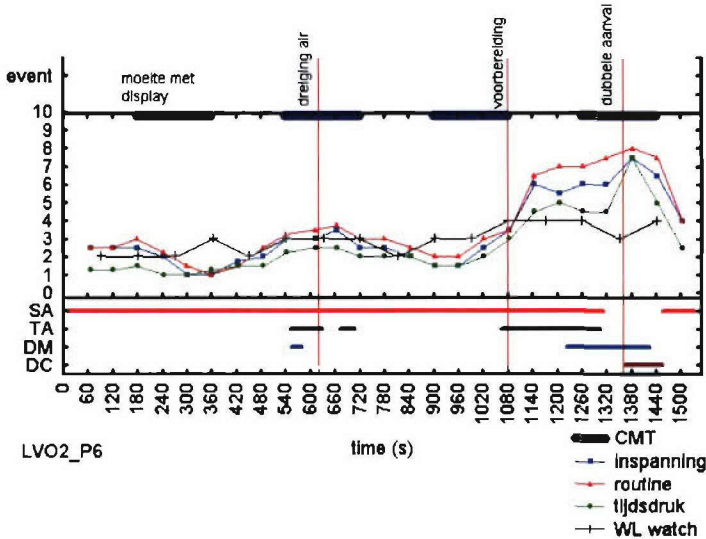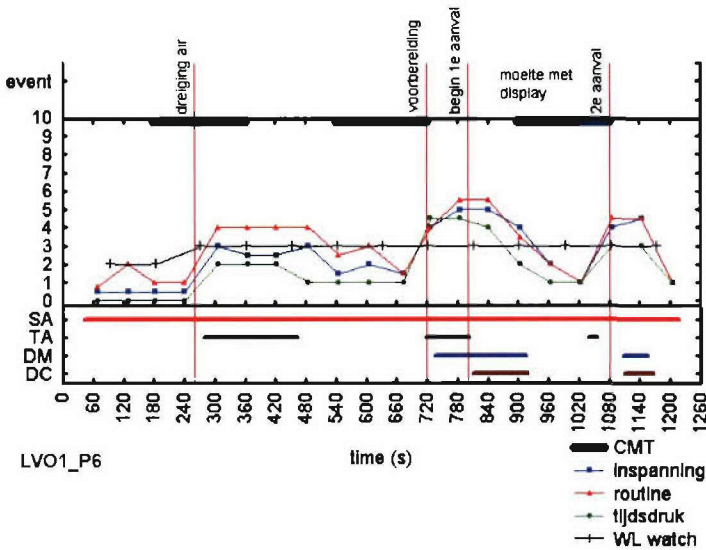| Signal (sample rate)	Sensor placement	Derived measures (unit)
Electro cardiogram (512 Hz)	Three electrodes on the chest	HR (beats / min) HRV mid-band (% from mean HR) HRV high-band (% from mean HR)
Electro oculogram (256 Hz)	Two electrodes above and below one eye	Blink frequency (blinks / min) Blink duration (ms) Blink amplitude (micro V)
Respiration (32 Hz)	Two inductive belts around the chest and abdomen	Respiratory frequency (Hz) Respiratory amplitude (arbitrary units)
Electro encephalogram (256 Hz)	Electrodes on the forehead and behind the right ear	(Not analyzed)

Data analysis

The physiological data were recorded with a Vitaport II system and were analysed with Vitagraph and SPIL (software tools belonging to the Vitaport system). The recorder was connected to the scenario computer in order to store relevant events. The start and the stop codes of the experimental sessions were used to synchronise the physiological data to the other data (such as the workload watch and subjective data of the video analysis).

Electrocardiogram (ECG)

The ECG was recorded with a sample rate of 512 Hz. The times at which R-peaks occurred were detected off-line by means of a SPIL algorithm. The results were inspected manually and artefacts were corrected or marked when a correction was not possible due to high noise levels. This occurred only a few times and only for a few seconds. Data during this periods were not included in the further analysis.

The time between to successive R-peaks were calculated and converted to HR values according to the procedure described by Velden and Graham (1988). The HR data was stored in a new channel of 4Hz.

The HR channel was used as input for the heart rate variability (HRV) analysis. A fast time frequency transform (F.T.F.T.) algorithm was used (Martens, 1992) to calculate variation in HR in two spectral ranges: the 'mid-frequency band' (0.075-0.15 Hz) and the 'high-frequency band (0.15-0.6 Hz). The results were stored in two channels of 1Hz each.

Respiration

Respiration was recorded with an additional Vitaport unit (PSG unit). This unit uses the 'inductive plethysmography' technique, which makes it possible to measure the surface within belts instead of the circumference, which is more common with other techniques. Two belts with coils were used (around the chest and the abdomen). The two signals were recorded at 32 Hz each. Respiratory frequency and amplitude were calculated by means of the F.T.F.T algorithm. The results were stored in two additional data channels

of 1 Hz each. Artefacts were identified by the phase shift between the chest and abdominal channel. A phase criterion of 60 degrees was used for artefact detection. These periods were excluded from further analysis.

Electro-oculogram (EOG)

The EOG was measured with two electrodes above and below one of the eyes. The EOG signals was stored with 256 Hz. Visual inspection of all signals showed no artefacts. Blink frequency and duration were calculated from the EOG signals with a SPIL algorithm. These parameters were stored in separate data channels of 4Hz. For the blink frequency the method op Velden and Graham (1988) was used to get proper frequency data within equal distance windows.

C   Description of the scenarios that were uses in the experiment

Table C.1   Description of the scenarios used in this study.

Scenario	Officer	Description
LVO1	ADO	After 3 minutes 4 air contacts head for the frigate from a SE direction. These contacts disappear after another 3 minutes. The officer knows they are in the vicinity. At t = 10:30 a decoy air contact performs a fly-by on the frigate, no attack is launched. After this, two groups of two air contacts perform a run on the frigate, in rapid succession. At t = 21:00 the scenario ends.
LVO2	ADO	This is a scenario full of decoys. It seems if some air contacts make a run on the frigate, but only in the end (at t = 22:00) an actual attack is made. Furthermore, the contacts behave in a non-routine manner. The officer needs to pay constant attention to the situation surrounding the frigate.
CCO1	PWO	In this scenario, the situation is clear. From the north, 3 surface contacts (boats) head for the frigate at top speed (37 kn). Since only an hostile boat is capable of this speed, countermeasures are soon deployed and the scenario ends around t = 15:00
CCO2	PWO	Four groups of surface contacts surround the frigate (complex surface picture). The officer needs to deploy the helicopter for reconnaissance of hostile movement. At t = 11:00, one group to the east shows hostile behaviour, and weapons are deployed. This scenario ends around t = 21:00
CCO3/ LVO3	PWO / ADO	In this scenario, one officer performs the tasks of two. At the start of the scenario, the picture is very complex and high in volume, both for air and for surface. The timing in this scenario is thus that around the end of the scenario (at t = 18:00), a simultaneous air and surface attack on the frigate is imminent. Before this, the officer has to be aware of all the contacts in the vicinity. In short, 2 air contacts threaten the frigate from the north-west, and 4 surface contacts from the east. Surrounding these hostile contacts are a lot of neutral contacts, rendering a complex scenario. End time is t = 21:00.

LVO = LuchtVerdedigingsOfficier (Principal Warfare Officer (PWO)

CCO = CommandoCentraleOfficier (Air Defence Officer (ADO)

Both the CCOs and LVOs participated in three different scenarios (see tabel). The third scenario was identical for the two groups (see [vanDelft et al., 2004] see for a more extensive description of the scenario and procedures).

Table C.2   Description of the scenarios used in this study.

Scenario	Officer	Description
LVO1	ADO	After 3 minutes 4 air contacts head for the frigate from a SE direction. These contacts disappear after another 3 minutes. The officer knows they are in the vicinity. At t = 10:30 a decoy air contact performs a fly-by on the frigate, no attack is launched. After this, two groups of two air contacts perform a run on the frigate, in rapid succession. At t = 21:00 the scenario ends.
LVO2	ADO	This is a scenario full of decoys. It seems if some air contacts make a run on the frigate, but only in the end (at t = 22:00) an actual attack is made. Furthermore, the contacts behave in a non-routine manner. The officer needs to pay constant attention to the situation surrounding the frigate.
CCO1	PWO	In this scenario, the situation is clear. From the north, 3 surface contacts (boats) head for the frigate at top speed (37 kn). Since only an hostile boat is capable of this speed, countermeasures are soon deployed and the scenario ends around t = 15:00
CCO2	PWO	Four groups of surface contacts surround the frigate (complex surface picture). The officer needs to deploy the helicopter for reconnaissance of hostile movement. At t = 11:00, one group to the east shows hostile behaviour, and weapons are deployed. This scenario ends around t = 21:00
CCO3/ LVO3	PWO / ADO	In this scenario, one officer performs the tasks of two. At the start of the scenario, the picture is very complex and high in volume, both for air and for surface. The timing in this scenario is thus that around the end of the scenario (at t = 18:00), a simultaneous air and surface attack on the frigate is imminent. Before this, the officer has to be aware of all the contacts in the vicinity. In short, 2 air contacts threaten the frigate from the north-west, and 4 surface contacts from the east. Surrounding these hostile contacts are a lot of neutral contacts, rendering a complex scenario. End time is t = 21:00.

LVO = LuchtVerdedigingsOfficier (Principal Warfare Officer (PWO))

CCO = CommandoCentraleOfficier (Air Defence Officer (ADO))

REPORT DOCUMENTATION PAGE
(MOD-NL)

1. DEFENCE REPORT NO (MOD-NL)	2. RECIPIENT'S ACCESSION NO	3. PERFORMING ORGANIZATION REPORT NO
TD2005-0450		TNO-DV3 2005 A245

4. PROJECT/TASK/WORK UNIT NO	5. CONTRACT NO	6. REPORT DATE
014.63063	-	January 2006

7. NUMBER OF PAGES	8. NUMBER OF REFERENCES	9. TYPE OF REPORT AND DATES COVERED
54 (incl 3 appendices, excl RDP & distribution list)	61	Final

10. TITLE AND SUBTITLE

The role of operator state assessment in adaptive automation

11. AUTHOR(S)

J.A. Veltman, M.Sc. and C. Jansen, M.Sc.

12. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

TNO Defence, Security and Safety, P.O. Box 23, 3769 ZG Soesterberg ,
Kampweg 5, Soesterberg, The Netherlands

13. SPONSORING AGENCY NAME(S) AND ADDRESS(ES)

TNO Defence, Security and Safety, Location Soesterberg, Human in Command, J.H. van Delft,
P.O. Box 23, 3769 ZG Soesterberg, Kampweg 5, Soesterberg, The Netherlands

14. SUPPLEMENTARY NOTES

The classification designation Ongerubriceerd is equivalent to Unclassified, Stg. Confidentieel is equivalent to Confidential and Stg. Geheim is equivalent to Secret..

15. ABSTRACT (MAXIMUM 200 WORDS (1044 BYTE))

Computer systems are capable to take over more and more tasks from human operators, but this does not always improve the performance of the human-machine system. Automation of which the level is made dependent on the situation may be an improvement for the performance. One of the parameters that might be used for this so-called adaptive automation is the state of the operator that can be estimated with physiological parameters.
This report provides a literature review on physiological measures and adaptive automation, and a model that describes the relation between the state of the operator, the human information processing and the interaction with the outside world. Furthermore, the results of a laboratory experiment are discussed. In this experiment, several physiological measures were monitored during the task performance of operators from the Netherlands Navy.
The literature review shows that physiological measures are promising for adaptive automation. The model is used to argue that there are many situations in which state estimators might not be useful for adaptive automation. Moreover, the results of the experiment show that it is difficult to measure the state within small time segments which is necessary for adaptive automation.

16. DESCRIPTORS	IDENTIFIERS
Mental workload	State assessment
Adaptive automation	Psychophysiology

17a.SECURITY CLASSIFICATION (OF REPORT)	17b.SECURITY CLASSIFICATION (OF PAGE)	17c.SECURITY CLASSIFICATION (OF ABSTRACT)
Ongerubriceerd	Ongerubriceerd	Ongerubriceerd

18. DISTRIBUTION AVAILABILITY STATEMENT	17d.SECURITY CLASSIFICATION (OF TITLES
Unlimited Distribution	Ongerubriceerd

Distributionlist

Onderstaande instanties/personen ontvangen het managementuittreksel en de distributielijst van het rapport.

4 ex.   DMO/SC-DR&D

1 ex.   DMO/ressort Zeesystemen

1 ex.   DMO/ressort Landsystemen

1 ex.   DMO/ressort Luchtsystemen

2 ex.   BS/DS/DOBBP/SCOB

1 ex.   MIVD/AAR/BMT

1 ex.   Staf CZSK

1 ex.   Staf CLAS

1 ex.   Staf CLSK

1 ex.   Staf KMar

1 ex.   TNO Defensie en Veiligheid, Algemeen Directeur,
ir. P.A.O.G. Korting

1 ex.   TNO Defensie en Veiligheid, Directie
Directeur Operaties, ir. C. Eberwijn

1 ex.   TNO Defensie en Veiligheid, Directie
Directeur Kennis, prof. dr. P. Werkhoven

1 ex.   TNO Defensie en Veiligheid, Directie
Directeur Markt, G.D. Klein Baltink

1 ex.   TNO Defensie en Veiligheid, vestiging Den Haag,
Manager Waarnemingssystemen (operaties), dr. M.W. Leeuw

1 ex.   TNO Defensie en Veiligheid, vestiging Den Haag,
Manager Beleidsstudies Operationele Analyse &
Informatie Voorziening (operaties), drs. T. de Groot

1 ex.   TNO Defensie en Veiligheid, vestiging Rijswijk,
Bescherming, Munitie en Wapens (operaties), ir. P.J.M. Elands

1 ex.   TNO Defensie en Veiligheid, vestiging Rijswijk,
Manager BC Bescherming (operaties), ir. R.J.A. Kersten

1 ex.   TNO Defensie en Veiligheid, vestiging Soesterberg,
Manager Human Factors (operaties), drs. H.J. Vink

Onderstaande instanties/personen ontvangen een volledig exemplaar van het rapport.

1	DMO/SC-DR&D standaard inclusief digitale versie bijgeleverd op cd-rom
2/3	DMO/DR&D/Kennistransfer
4	Koninklijke Marine, DMKM/CAMS-Force Vision KLTZ J.J. Teertstra, Programmabegeleider Defensie en projectbegeleider
5	Projectleider TNO Defensie en Veiligheid, Soesterberg, Business Unit Human Factors, Drs. J.A. Veltman
6/8	Bibliotheek KMA
9	Koninklijke Biblitheek
10	Programmaleider TNO Defensie en Veiligheid Business Unit Human Factors, Drs. J.H. van Delft
11/12	TNO Defensie en Veiligheid, vestiging Soesterberg, Archief
13	TNO Defensie en Veiligheid, Soesterberg, Business Unit Human Factors, Drs. J.A. Veltman
14	TNO Defensie en Veiligheid, Soesterberg, Business Unit Human Factors, Drs. C. Jansen
15	TNO Defensie en Veiligheid, Soesterberg, Business Unit Human Factors, dr. H.A.H.C. van Veen
16	TNO Defensie en Veiligheid, vestiging Soesterberg, Marketing en Communicatie, digitale versie via Archief
17	TNO Defensie en Veiligheid, vestiging Rijswijk, afdeling Tekstverwerking (digitale versie)